LACIE-00712
JSC-10819

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

7.9-10140
CR-158149

NASA NOAA USDA

(E79-10140) LARGE AREA CROP INVENTORY EXPERIMENT (LACIE). CAMS DETAILED ANALYSIS PROCEDURES (Lockheed Electronics Co.) 159 p HC A08/MF A01 CSCL 02C

N79-18405

Unclas
G3/43 00140

CAMS DETAILED ANALYSIS PROCEDURES

National Aeronautics and Space Administration

LYNDON B. JOHNSON SPACE CENTER

Houston, Texas

March, 1976

CLASSIFICATION AND MENSURATION SUBSYSTEM (CAMS) DETAILED ANALYSIS PROCEDURES

APPROVED BY:

F. G. Hall
CAMS Subsystem Manager

R. M. Bizzell
CAMS Operations Manager

MARCH 1976

ACKNOWLEDGEMENTS

This document represents the work of personnel from Lockheed Electronics Company, Inc./Aerospace Systems Division, the National Aeronautics and Space Administration, and the U. S. Department of Agriculture. This work was performed under NASA Contract NAS 9-12200.

CONTENTS

Section Page

Section | Page

Section | Page

Appendix Page

FIGURES

ACRONYMS

AI	Analyst Interpreter
B/W	Black and White
C&I	Composition and Indexing
CAMS	Classification and Mensuration Subsystem
CAS	Crop Assessment Subsystem
CEF	CAMS Evaluation Form
CIR	Color Infrared
COM	Computer Output Microfiche
CRD	Crop Reporting District
DAPTS	Data Acquisition, Preprocessing, and Transmission Subsystem
DEAF	Dell Foster Deck
DM	Data Manager
DO	Designated Other
DPA	Data Processing Analyst
DPCA	Data Production Control Analyst
DPR	Data Product Request
DR	Discrepancy Report
DU	Designated Unidentifiable
ERIPS	Earth Resources Interactive Processing System
ERTS	Earth Resources Technology Satellite (renamed Landsat)
FLAP	Final Product
FLOCON	Flow Control
GSFC	Goddard Space Flight Center
ID	Identification

Figure | Page

LACIE	Large Area Crop Inventory Experiment
Landsat	Land Satellite (formerly ERTS)
LARS	Laboratory for Applications of Remote Sensing
LDSS	LACIE Data Systems Supervisor
LPDL	LACIE Physical Data Library
MLA	Mean Level Adjustment
MSS	Multispectral Scanner
OCC	Operations Control Center
PFC	Production Film Converter
PTD	Photographic Technology Division
USDA	United States Department of Agriculture

1. INTRODUCTION

The detailed processing procedures for a unified team approach for estimating wheat acreages from the Land Satellite (Landsat) Multispectral Scanner (MSS) data are presented in this document. Analyst Teams, consisting of two analysts assisted by a consulting Regional Analyst, will process the Large Area Crop Inventory Experiment (LACIE) sample segment data. The analysts will perform all processing functions [i.e., training field selection, interactive reworks on the Earth Resources Interactive Processing System (ERIPS) and classification results evaluation] as a team.

The major emphasis during the second phase of LACIE is on the identification of wheat and the estimation of the acreage harvested. The procedures that follow reflect concentration on this single crop and lesser emphasis (therefore smaller samplings) on other crop and non-crop categories.

The Analysts will select various training fields, exclusion areas, and test fields from one or more acquisitions collected during the crop year. These fields will be extracted from 5-mile by 6-mile segments dispersed throughout major wheat growing areas and utilized in automatic data processing classification programs to determine total wheat acreage for these segments.

During LACIE Phase II, the Analyst Interpreters (AI) and Data Processing Analysts (DPA) will interpret and process LACIE segments in a unified team approach. Considerable assistance will come from the Regional Analysts who will provide inputs to the Analyst Team during the review process, the interpretation phase and during the segment evaluation subsequent to batch or interactive processing. Each Regional Analyst will support several Analyst Teams.

Segments will be assigned to the teams based on the priority of the segment and the regional emphasis of the team. Segments which have been processed in the same biowindow with a satisfactory evaluation rating will not be assigned again until the next biowindow.[1] Those segments which previously received a marginal or unsatisfactory evaluation rating in the same biowindow as the current acquisition will be reviewed according to normal procedure and reprocessed.

This procedures document describes in detail the functions depicted in the *Classification and Mensuration Subsystem (CAMS) Functional Procedures Document*.

---

[1]An exception is biowindow 1 for winter wheat. Subsequent acquisition may be reviewed before the opening of biowindow 2.

2. IMAGE INTERPRETATION PROCEDURES

The primary function of image interpretation in LACIE is to identify and select wheat and nonwheat training fields for training the pattern recognition algorithm to classify all pixels within the designated segment.

2.1 IMAGERY REVIEW

An overview of the review procedures and the purpose of the Analyst Status Form are described below.

a. Review Prior to Initial Training Field Selection

During Phase II, all segments are designated as training segments and will be acquired at every opportunity, except during periods of snow cover.

When an acquisition is beyond the Robertson phenological stage of 2.3, according to the adjusted crop calendar, an estimate of the wheat[1] or small grains present in the segment will be made. If the acquisition is prior to stage 2.3 and an estimate is not attempted, the Data Acquisition/Evaluation Record is completed by indicating whether the acquisition was unworkable due to pre-emergence, clouds, haze, or snow and is passed to the Crop Assessment Subsystem (CAS). The Regional Analyst will assist the Team Analysts on a consultation basis by providing real-time updates to ancillary data and crop calendars. (NOTE: Training fields may be selected from acquisitions which had been previously reviewed with no training fields selected. This situation may occur when the proper signature for wheat in a particular geographic region is not recognized early in the crop growing year. A wheat estimate, performed by manually

counting pixels on a low wheat segment, may also occur for the same reasons given above.)

[1]Wheat should be interpreted as "wheat category", i.e., wheat/small grains, unless otherwise noted.

b. Review Following Initial Training Field Selection

The Regional Analyst will participate during this process performing the same function as described in the preceding paragraph.

On subsequent acquisitions except those for which a satisfactory estimate has been given, all training fields will be checked to verify previous interpretations, and areas classified as wheat will be checked for changes such as winterkill. The review image will be compared to the classification map to determine if a better wheat estimate is possible using this data.

If small grains were identified on the initial interpretation, the review image should be interpreted to see if wheat can be separated from the other small grains; or further separated, if wheat and some other small grain(s) has been previously separated from all other small grains.

If there is sufficient change to warrant a revision of the previous estimate, training fields will be revised according to the procedure in section 2.4.

New training fields need not be selected for acquisitions following the initial training field selection if there has been less than 500 pixels changed.

c. Analyst Status Form

A record of the pertinent information related to all sample segments interpreted by an Analyst is kept on a sample segment status form (see fig. 2-1). The form contains the segment number and the county or district. In addition, calendar date and the date the Analyst begins and terminates the interpretation task are recorded. Additional data flow tracking information is included.

ANALYST STATUS FORM

Segment Number	Segment Name	Calendar Date	Date In	Date Out	Review Prior Training Field Selection	Review Following Training Field Selection	Training Field Selection	Comments

Figure 2-1. — Example of the Analyst's sample segment status form.

2.1.1 CHECKLIST OF SEGMENT PACKET CONTENTS

The input products received by the Analyst in the sample segment packet from the LACIE Physical Data Library (LPDL) are as follows:

a. Detailed Crop Calendar

b. Ancillary Summary

c. Topographic map(s) (large scale when possible)

d. Film Products

   First acquisition in biowindow:

   (1) False color infrared (CIR) composite (Bands 4, 5, and 7) — Product 1

   (2) Enhanced image (Bands 5, 6, and 7) — Product 2

   (3) Black and white (B/W) positive transparencies of Bands 4, 5, 6, and 7 — Product 4

   (4) All other acquisitions: Products 1 and 2 only

e. Previous Classification, If Applicable

   (1) Temporal Crop Interpretation Forms

   (2) CEF

   (3) Classification Map

   (4) Classification Results

   (5) Primary Overlays

   If any of the materials from a, b, c, or d are missing, the Analyst will submit a discrepancy report (DR) to the Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS) requesting the required data.

2.1.2 QUALITY EVALUATION OF CONTENTS

The evaluation criteria for imagery quality are detailed in appendix A.

2.1.3 PREPARATION OF IMAGERY

All Product 1's will be annotated with the segment number, county name, date of acquisition, and pixel/line numbers. NOTE: The segment number will be identified with the segment number, followed by the biostage and the Julian date. An example is given below:

$$\underbrace{1960}_{\text{Segment number}} - \underbrace{2}_{\text{Biostage}} - \underbrace{131 - 6}_{\text{Julian date}}$$

When the above task is completed, an overlay will be prepared and registered to the image for annotations. Field boundaries, identifications, and other pertinent data will be recorded on this overlay for reference during the review of the acquisition.

2.1.4 LOCATION OF SEGMENT

The segment will be accepted as received from Goddard Space Flight Center (GSFC) if it has common area with the predicted sample segment area. For those segments which are acceptable but whose center point falls outside the intended political subdivision, a note of this fact is made on the CEF. In all other cases, the segment data will be rejected on the basis of segment mislocation. Those segments which deviated from the predicted location in the prior year, but were acceptable under the criteria at that time and do not meet the above criteria, will be referred to CAS for a decision on acceptance or rejection. The procedure for locating the segment is described below.

The segment boundaries are plotted onto the large scale map provided in the Sample Segment Packet. The plotting procedure begins with the Analyst locating the center point of the sample segment on the map. (This task was completed for many sample segments during LACIE Phase I.)

The geographic coordinates for the segment are extracted from the ancillary summary in the Team Packet. When the center point has been located and plotted onto the map, a segment boundary overlay is positioned over this point. The Analyst positions the overlay at the appropriate angle to compensate for the attitude of the satellite ground track. After this is accomplished, the Analyst examines the sample segment imagery for natural or cultural features in order to precisely locate the boundaries of the segment on the map. The county or political division boundaries are delineated on the map in order to assist the Analyst in interpreting the county statistics.

Sample segments within the United States may be located on the United States Department of Agriculture (USDA) Soil Conservation Service Earth Resources Technology Satellite (ERTS) 1 mosaics (scale 1:1,000,000) if full frames are not available. In addition, the following data and procedures can be used to locate and verify new and existing LACIE sample segments.

2.1.4.1 Landsat Full Frames

Landsat 1 or 2 full-frame CIR imagery covering each segment will be available for locating selected segments. The procedures for using the full frames are:

a. The analyst will locate the sample segment on the Landsat full frame. The outline of the segment will be plotted onto a clear overlay keyed to the 9.5-in. color positive. The large area coverage (100 n. mi. × 100 n. mi.) will allow

the interpreter to examine the area surrounding the sample segment being interpreted.

b. Field patterns and signatures outside of the sample segment can be compared to those within it to aid the analyst in identifying wheat and other crops. Soil and soil moisture patterns that may affect signatures within the sample segment are often identifiable on the 9.5-in. full frame. For example, the analyst will be able to identify streams and sand dunes more easily on the full frame.

c. The full frame will be used to precisely locate all sample segments because many topographic maps do not contain sufficient detail for precisely locating the segment.

d. Large area coverage also permits the analyst to examine all or part of the region or political division within which the sample segment is located. This is important because the agricultural statistics and crop calendar information are based on regions or political divisions (for example, crop reporting districts).

e. The large area coverage will be used to establish the "green wave" as the growing season progresses. This imagery will be reviewed, as it is acquired, as a contiguous strip of imagery overlayed on 1:1,000,000 scale maps to correlate imagery "green wave" with latitude and longitude. This information will be a check against statistical models which predict wheat biophases.

2.1.4.2 Use of 1974-1975 Acquisitions

Acquisitions from the 1974-1975 growing season will be available for verifying existing segment locations. After the segment is plotted for the first time, the location of all subsequent acquisitions will be verified and replotted if necessary.

2.1.5 FAMILIARIZATION WITH BACKGROUND DATA

The Analyst will study the crop calendar, *Weekly Weather and Crop Bulletins*, historical data, topographic map(s), ERTS mosaics, Landsat-2 full frames, and the *Wheat Identification Aid* pertaining to the sample segment. When this procedure is complete, the Analyst begins the review for training fields (sections 2.1.7 and 2.1.8).

2.1.6 REVIEW FOR DETERMINATION OF WHEAT ESTIMATE

a. Review the adjusted crop calendar and meteorology data for the current acquisition(s).

b. Examine the imagery for emerging wheat signatures and field patterns.

c. If the previous year's imagery is available for temporal interpretation, locate the potential small grain fields, if applicable, and relate to the current imagery and any earlier acquisitions.

d. When the previous year's imagery is not available, interpret the current imagery and all earlier acquisitions.

e. Identify field patterns and signatures and note the changes in the development patterns.

f. If only a single acquisition is available for review, consult the Regional Analyst for additional historical and/ or current information on this segment.

g. If the adjusted crop calendar indicates a biological stage of less than 2.3 or greater than or equal to 9.0, an estimate will not be made and the appropriate code passed to CAS.

h. The analyst will determine the current biowindow according to procedures in section 2.3.1.4.5.

i. If the wheat is in a growth stage of greater than 2.3, an estimate will be made and the growth stage entered on the CEF.

j. Record the necessary information, as described in section 3.4.6.3, on the Data Acq./Eval. Record. If an estimate is not made, no CEF is required.

2.1.7 REVIEW FOR SIGNIFICANT CHANGE IN WHEAT ESTIMATE

The purpose of this review is to determine if the wheat estimate has significantly changed on the most recent acquisition. The procedures are as follows:

a. The current acquisition is compared to the classification map and Product 1 from the previous wheat estimate.

b. The Analyst interprets for any changes in the spatial arrangement of wheat.

c. The Analyst must determine if any previous wheat and non-wheat areas are now identified as nonwheat or wheat on the current acquisition.

d. The determination of additional wheat fields must also be made.

e. If the change in the wheat estimate is less than 500 pixels, CAS is informed by code via the Daily Report to CAS. The analyst completes the Data Acq/Eval. Record. No CEF is necessary.

f. When the wheat estimate change is greater than an absolute 500 pixels, the Analyst must revise the estimate.

g. The processing modes for revising a wheat estimate are:

(1) Select new training fields for batch processing.

(2) Interactively revise training fields on ERIPS.

(3) Estimate the wheat change by overlaying a 250-random-dot overlay onto the classification map and compute the

percentage of change. The areas of change will be determined by interpreting the imagery and delineating the areas of change on an overlay. The random-dot overlay and wheat change overlay are registered to the Product 1. The new estimate is computed by dividing the number of dots within the wheat change areas by the total number of random dots (250).

2.1.8 CLUSTER MAP GENERATION

The Analyst *may* choose to have the segment clustered prior to identifying signatures. *If* a cluster map is desired, it may be generated batch or interactive, depending on the availability of time. See appendix E for detailed instructions.

2.2 PROCESSING MODE SELECTION

This section describes the criteria for: (1) Selecting the best acquisition for estimating wheat when multiple acquisitions are in the team packet, (2) manually counting the amount of wheat on a segment with a low proportion of wheat, and (3) selecting multitemporal signatures for computer classification.

2.2.1 ACQUISITION SELECTION AND PROCESSING MODE

When multiple acquisitions arrive in the team packet for analysis, the following procedures must be used for selecting the proper acquisition for processing.

a. When the estimate of wheat is the same on all acquisitions, select training fields or count wheat on the *first* acquisition.

b. If there are significant differences in the wheat estimate among the various acquisitions, use the most recent acquisition with the significant change.

c. When there is no confusion between the wheat and nonwheat, the chosen acquisition will be processed.

d. If confusion exists between the wheat and non-wheat, the confusing nonwheat areas should be designated other (DO).

e. When DO is impractical due to many small confusing nonwheat areas, multitemporal processing should be attempted.

f. When the confused areas are within the system tolerances (±500 pixels) the acquisition should be processed even if the confusion cannot be eliminated by DO or multitemporal processing. All comments should be noted on the CEF.

2.2.2 LOW WHEAT SEGMENTS

Prior to selecting wheat and nonwheat training fields it must be determined if the number of wheat pixels exceed approximately 1200 pixels or 5 percent of the segment. When the amount of wheat is less than 1200 pixels, all pixels of wheat will be counted manually. The fields containing wheat will be delineated and annotated on the primary overlay. The number of pixels within each field is recorded on the primary overlay adjacent to the field boundary.

The percentage of wheat is computed by dividing the total number of pixels in the segment (22,932) into the total number pixels of wheat in the segment. The wheat percentage will be recorded on the CEF (fig. 2-2) and forwarded to CAS.

One copy of the CEF will be forwarded to CAS, one copy to the Team Envelope, and one copy to the CAMS Control Clerk.

Other signatures and identifications should be noted for reference during the review of subsequent acquisitions.

CAMS EVALUATION FORM

SEG# ☐☐☐☐ ☐ ☐☐☐ ☐

RESULTS ________ __/__/__

ANALYST(S) ____________________

DPR # ____________________

LOCATION ____________________

☐ PRE-EMERGENCE

☐ CLOUDS/HAZE/SNOW

☐ NO SIGNIFICANT CHANGE

☐ LESS THAN 5% WHEAT/ SMALL GRAINS ________ PIXELS

AI BIOPHASE ____________________

NW IDENTIFICATION:

ACQUISITION DATES

____

____

____

____

☐ FEATURE SELECTION ____________________

☐ STATISTICS MANIPULATION ____________________

ACCURACY

			TEST			TRAINING		
			%	CAT	%TH	%	%TH	NAME
W	______							
W	______							
W	______							
W	______	P ____	____	____	____	W ____	____	____
N	______	P ____	____	____	____	W ____	____	____
DO	______	P ____	____	____	____	W ____	____	____
THRSH	______	P ____	____	____	____	W ____	____	____
X*X	______	P ____	____	____	____	N ____	____	
DU	______							

Evaluation Form Revised 12/12/75

Figure 2-2. – CAMS Evaluation Form.

2.2.3 MULTITEMPORAL CLASSIFICATION DECISION

a. Conditions Necessary Prior to Decision

(1) The decision has been made to change the wheat estimate by selecting training fields on a chosen acquisition and classifying.

(2) "Confusion" exists between wheat and nonwheat on the chosen acquisition.

(3) There are one or more other acquisitions available which are registered to the chosen acquisition.

b. Multitemporal Decision

If confusion exists between wheat and nonwheat on the single acquisition and another acquisition exists that eliminates or reduces the confusion areas, then a multitemporal interpretation and classification will be made on the segment. In general, if the interpreter requires the additional imagery to determine wheat field identification, the same information will be supplied to the classifier.

c. Conditions that Inhibit Multitemporal Processing

(1) Misregistration (see appendix A)

(2) Clouds, cloud shadows, or haze on any acquisition considered such that the total covered area (all acquisitions considered) fails to meet the criteria in appendix A

(3) Subclass sample size too small ($n + 1$, where n is the number of channels)

(4) Number of subclasses exceed software limitations (60)

2.3 TRAINING FIELD SELECTION

Wheat and nonwheat training fields will be selected for two basic types of computer classification. First, single acquisition classification entails selecting signatures from a single Landsat acquisition. The interpretation of the signatures may require the Analyst to temporally interpret multiple acquisitions,

but the computer classification will be accomplished by using four channels from one acquisition. The second multitemporal classification requires signatures from more than four channels of data. Multitemporal classification will normally involve eight channels of Landsat MSS data.

The procedures for the training field selection techniques described above are detailed in the following sections.

2.3.1 SINGLE ACQUISITION TRAINING FIELD SELECTION

The procedures in this section describe the selection of training fields from a single acquisition. The procedures for selecting multitemporal signatures and training fields are detailed in section 2.3.2.

2.3.1.1 Analyst Decision Criteria

A portion of the decision criteria used by the Analyst during the training field selection and review processes is shown in figure 2-3.

The decision tree is not complete and will be updated on a continuing basis throughout Phase II.

The types of decisions shown in figure 2-3 are used to interpret the LACIE sample segment for wheat and non-wheat training fields.

2.3.1.2 Imagery Overlay Preparation

All field corners and field codes will be recorded on the primary overlay.

2.3.1.3 Exclusion Areas Definition

The detailed procedures for excluding areas from computer classification are in section 2.3.1.5.3.

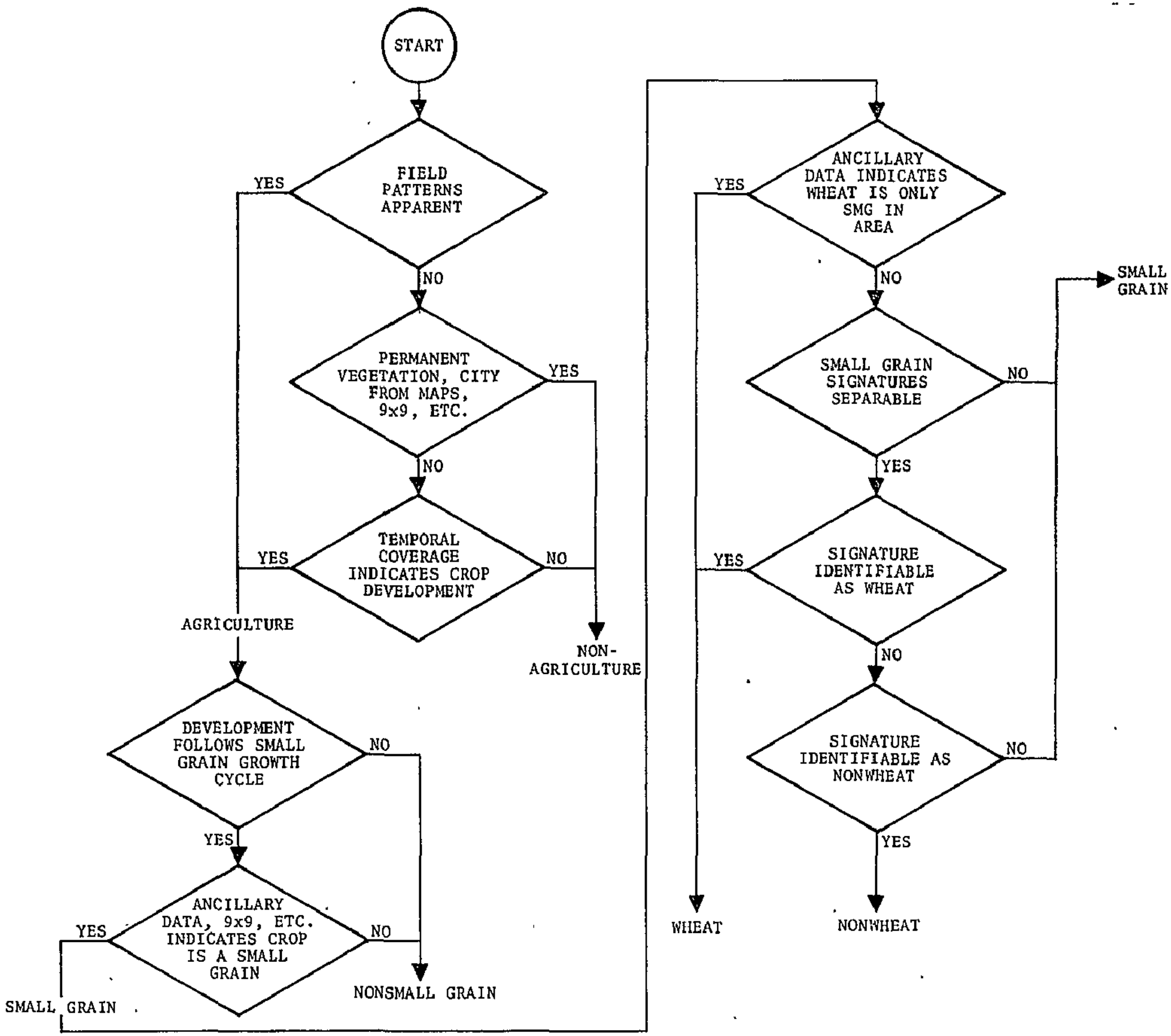


Figure 2-3. — A portion of the decision criteria used by the Analyst.

The first task in selecting fields is to designate areas that will be excluded from the LACIE automatic classification process, thus avoiding possible signature confusion. The two types of exclusion areas are DO and Designated Unidentifiable (DU). See appendix B for coding instructions. The DO areas will be excluded from the classification process by using the 1974-1975 growing season sample segments and Landsat 2 full frames as an aid in interpreting the LACIE Phase II imagery. (See section 2.1.4.1 for a more detailed explanation of the use of Landsat full frames for excluding areas from computer classification.) The DO areas delineated on the 1974-1975 imagery should be reviewed to determine whether they can be used to aid interpretation of the current DO areas.

The two types of exclusion areas are defined below.

- DO — Field Type D
    a. Areas that are identified as nonwheat.
    b. Examples: Lakes, swamps, forests, mountains, streams, cities, airports, and any nonwheat area.
    c. Annotated by a set of 2-to-10 vertices. More than one set of vertices may be used, if required. It is recommended that at least three vertices be used.
    d. This type of designation is most useful when areas have similar signatures to wheat, yet can be identified as nonwheat. On certain dates, for example, wheat fields and forested areas have similar signatures. Transmitting the forested area as DO will eliminate it from classification and thus prevent misclassification of the wheat fields.
- DU — Field Type U
    a. Areas that are not identifiable and thus could possibly contain wheat (ground is obscured). This classification should be used for areas affected by

haze, even though field signatures are discernible through the haze. It is recommended that this procedure be used even though training fields may be selected in DU areas (see section 2.3.1.5.3, item b).

b. Examples: Clouds, cloud shadows, flooded areas, snow-covered areas, haze-covered areas, etc.

c. Annotated by a set of 2-to-10 vertices. More than one set of vertices may be used, if required. It is recommended that at least three vertices be used.

2.3.1.4 Training Field Identification Criteria

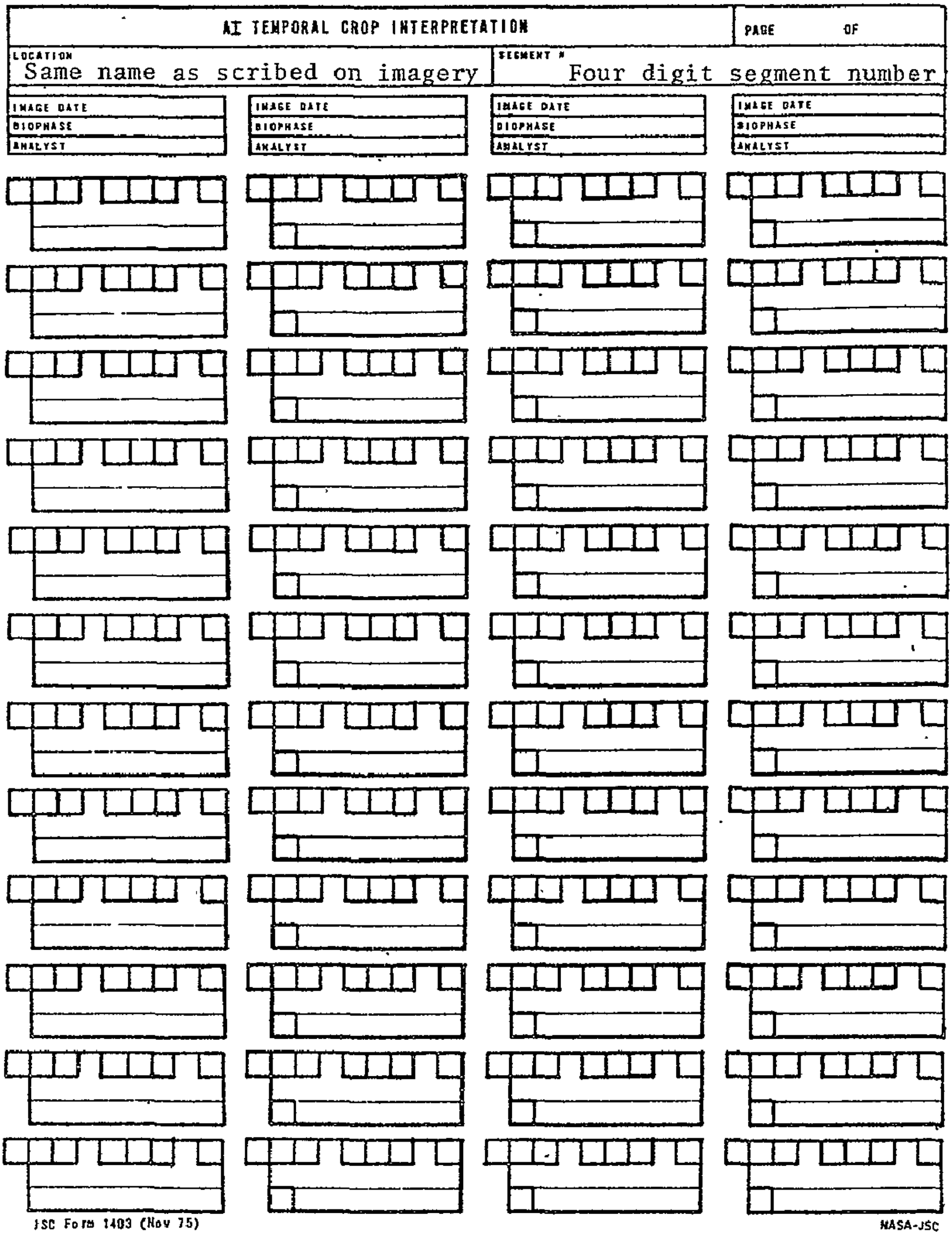
All field identifications will be recorded on the AI Temporal Crop Interpretation Form (fig. 2-4). On this form are four vertical columns of "field identifier elements" (One for each Data Base Update) which are designed to contain all the identification necessary for the fields data base. This information is used at the Dell Foster digitizer (see appendix D) when creating field coordinates. Information should be recorded on the form in as much detail as possible with the first interpretation, since the "no change" code and other codes can be used in subsequent columns, reducing the amount of work at the Dell Foster. See pages 2-18a thru 2-18c for instructions. Further information on coding standards is contained in appendix B.

Several criteria must be followed before training fields are acceptable for classification. The sections that follow define the minimum field and category sizes, the use of the cluster map, the use of the crop calender, and the use of crop reporting unit statistics from the ancillary summary.

2.3.1.4.1 Signature Identification

The identification (ID) of specific signatures is accomplished through the use of multiple acquisitions which allow the Analyst to trace the history of a field or feature through one or more cropping seasons. Temporal analysis of the Landsat MSS data permits the Analyst to associate certain signatures with various ground conditions such as bare soil, water, pastures, and green crops.

The signatures for wheat, other small grains, and nonwheat are illustrated in the *Wheat Identification Aid*, which has temporal Landsat MSS data of several intensive test sites in the United States and Canada. The signatures of these crops are representative of this region only and caution should be used when extending these signatures beyond the regions they represent.



AI TEMPORAL CROP INTERPRETATION

PAGE OF

LOCATION: Same name as scribed on imagery

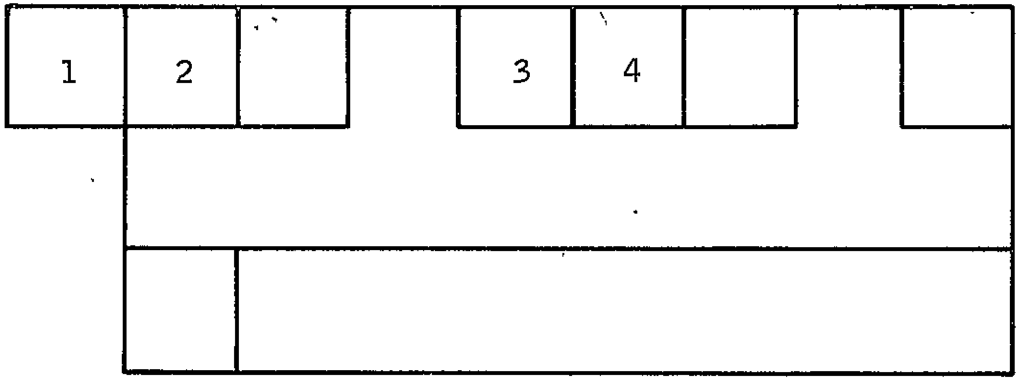
SEGMENT #: Four digit segment number

IMAGE DATE	IMAGE DATE	IMAGE DATE	IMAGE DATE
BIOPHASE	BIOPHASE	BIOPHASE	BIOPHASE
ANALYST	ANALYST	ANALYST	ANALYST

JSC Form 1403 (Nov 75)

NASA-JSC

Figure 2-4. — Temporal crop interpretation form.



1. Category

Wheat-W

Nonwheat-N

Unidentifiable-X

Test-P

Designated Other-D

Designated Unidentifiable-U

2. Field number

01→99
0A→9Z

3. Class

Winter Wheat-W
Spring Wheat-S
Small Grain-K
Winter Small Grains-B
Spring Small Grains-D

Agriculture-A
Nonagriculture-N

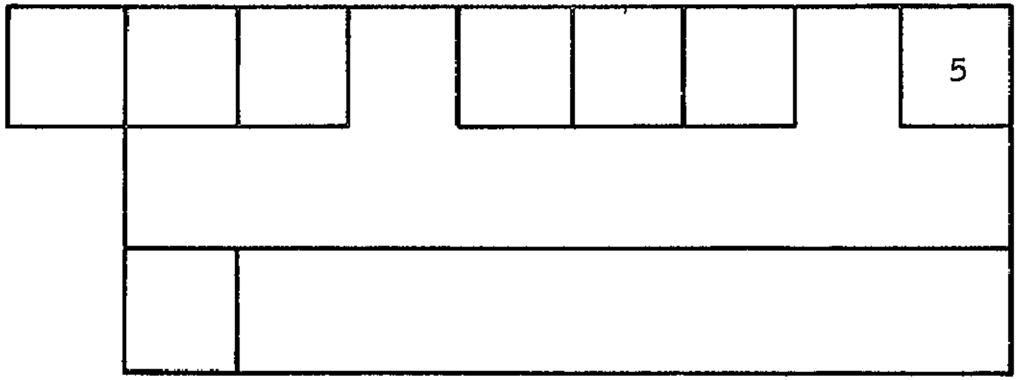
Unidentifiable-X

4. Subclass

1→50

Note: The field identifier elements are arranged in four vertical columns on the form - one for each Data Base Update. Do not use more than one column per Data Base Update.

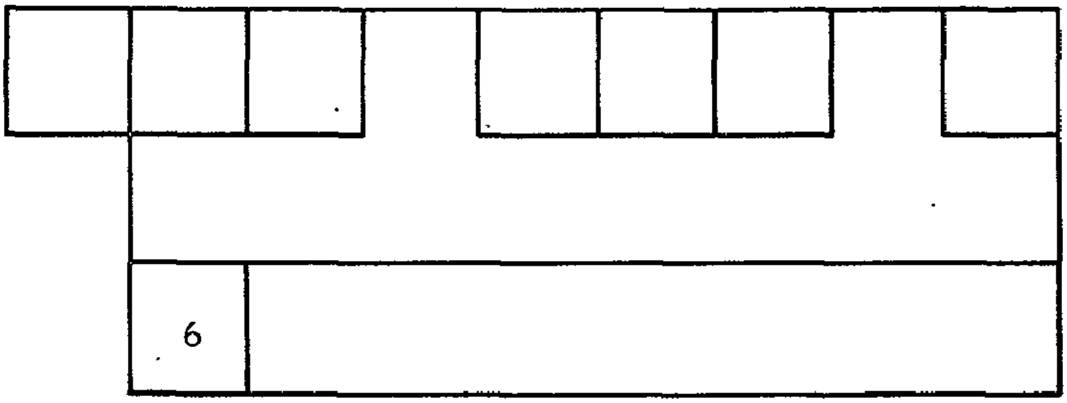
Field identifier element.



This box has two (2) purposes:

1) It was used with Signature Extension to create Mean Level Adjustment (MLA) fields. By placing an -A- in this box of a wheat category training field, the field would be duplicated but be placed in the MLA category (A) (i.e., when the coordinates for W03 were read the "A" was duplicated to A03 by LARS, thus saving time by not having to read the coordinates for this field. Signature Extension is not currently operational.

2) It is currently being used to allow retention of field coordinates in the data base when only a change in the class or subclass is desired. This is done by placing a -B- in the box. This is used <u>ONLY</u> when the only change is the class or subclass. If a change occurs in a training field's coordinates or its category, the appropriate change must be made on the Temporal Crop Interpretation Form and the coordinates must be reread.



Change codes (See appendix B, table B-1).

Signature identification keys for all LACIE wheat-growing regions are in progress and when the keys are completed they will become available to the Analyst.

The use of previously identified signatures during the Analyst's decision of wheat/nonwheat is briefly described in section 2.3.1.1.

Cluster map usage - The cluster map *may* be used as an aid during the interpretation process. The primary uses are to determine the minimum number of color classes the Analyst can expect and to locate the spatial distribution of the clusters within the segment.

2.3.1.4.2 Minimum Field Size

The recommended minimum field size is 25 pixels. This restriction is a recommendation, not an absolute minimum. Fields containing fewer pixels can be used when no other fields are available for training purposes. However, the analyst should be aware that small fields may cause interpretation and registration problems.

The minimum width is four pixel columns, and the minimum depth is one pixel line.

A minimum of 2 and a maximum of 10 vertices may be used. Training fields with two vertices (a straight line) cannot be oriented vertically paralleling pixel lines. They must be offset at a sufficient angle to meet the minimum width (four pixels) criterion. Figure 2-5 illustrates this restriction.

2.3.1.4.3 Minimum Category Size

The minimum number of pixels recommended for computer classification is 500 training pixels for wheat and 100 training pixels for each nonwheat subclass.

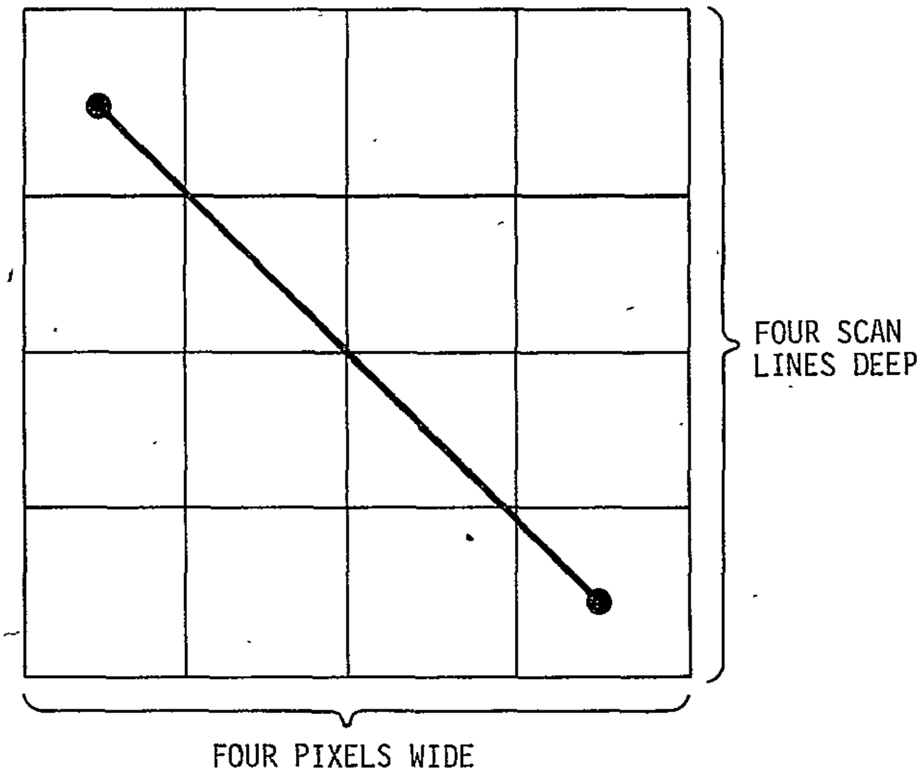


Figure 2-5. — Example of two vertice field
that is four pixels wide.

When the recommended minimums cannot be attained, the Analyst should select as many pixels of wheat and nonwheat as possible.

2.3.1.4.4 Use of Nominal Crop Calender

The nominal crop calenders for wheat and other major crops within the same crop reporting unit are displayed in graphic form and an example of a typical crop calender is shown in figure 2-6.

2.3.1.4.5 Use of adjustable crop calendar

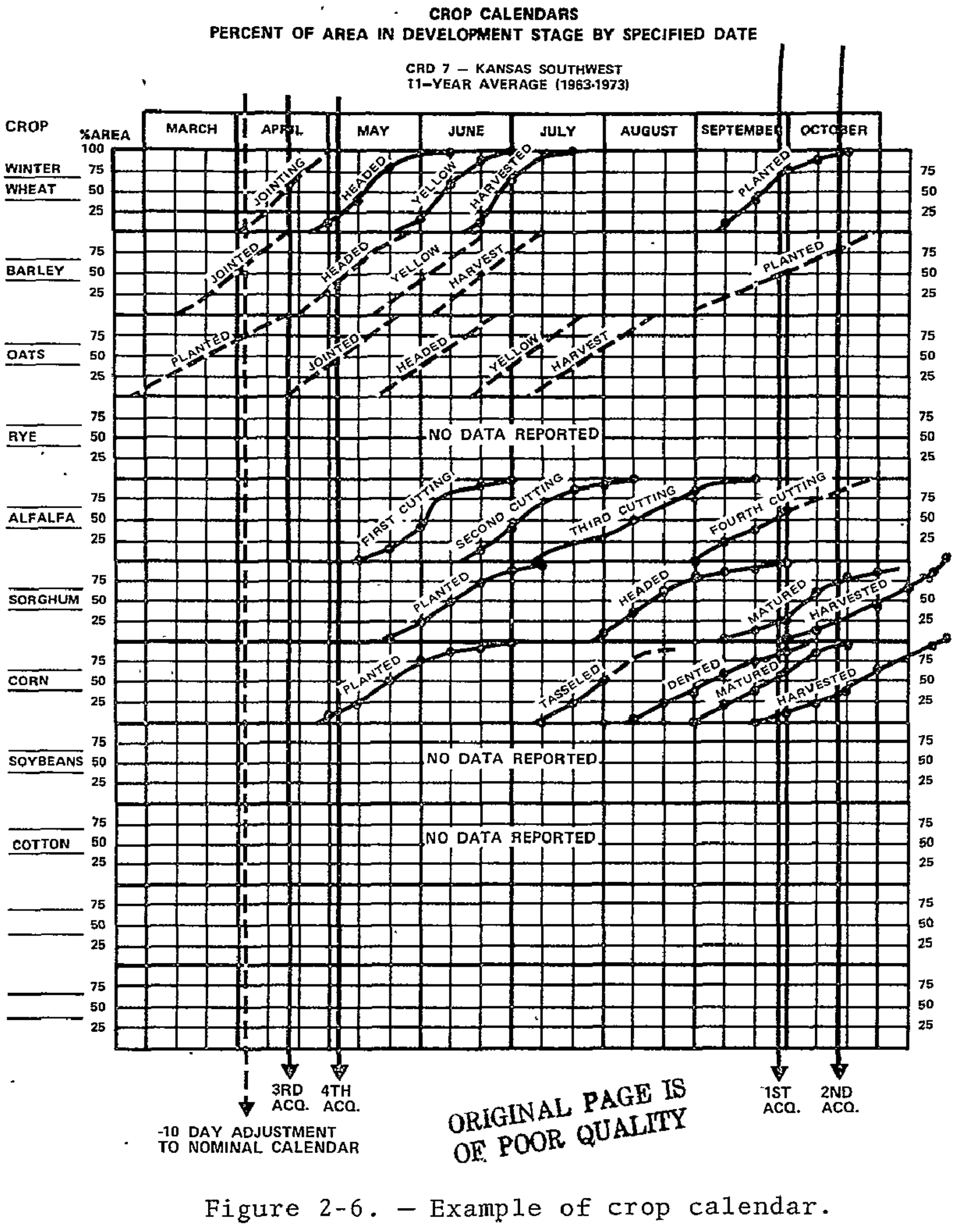
The adjustment in "plus or minus days" (delta days) is extracted from the *Weekly Weather and Crop Summary* and applied to the crop calendar before working the segment (fig. 2-6).

Figure 2-7 is an example of the data in the *Weekly Weather and Crop Summary* and illustrates the procedure for computing the number of days departure from the normal for individual crop reporting districts. The number in parentheses indicates the number of days the wheat development stage has departed from the norm. The number "-11.5" for crop reporting district (CRD) 1 shows that the development stage is 11.5 days behind the normal development at that calendar date. The number "5.0" in CRD 1 indicates that 50 percent of the wheat is in the soft dough stage and is additional information for use by the Analyst.

The procedures for adjusting the nominal crop calendar are given below.

a. Draw a vertical line on the crop calendar through the wheat biological stage during which the imagery data set was collected (fig. 2-6).

b. If the wheat biological stage differs from other crops and a color class can be identified as wheat (*Wheat Identification Aid* can be used as a guide to color trends of wheat in



ORIGINAL PAGE IS OF POOR QUALITY

Figure 2-6. — Example of crop calendar.

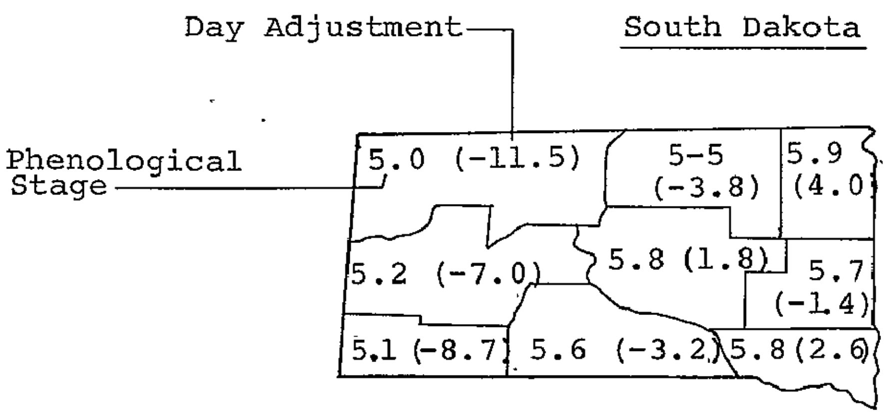


Figure 2-7. — Example of crop calendar adjustment in the Weekly Weather and Crop Summary.

various biological stages), designate approximately 500 pixels of wheat for training.

c. If the imagery does not support the crop calendar, estimate the actual growth stage of wheat.

d. If, through adjustments and/or the Analyst's estimates, the biowindow changes, record this on the CEF to notify CAS of the proper designation.

The biowindow changes if the majority of the fields on the segment are in a growth stage other than the one specified on the data base.

The biowindows are defined as follows:

Biowindow 1 - Crop Establishment - From emergence to jointing

Biowindow 2 - Greening Up - From beginning of jointing to end of booting

Biowindow 3 - Heading - From start of heading to soft dough

Biowindow 4 - Ripening - From soft dough to harvest.

2.3.1.4.6 Use of Crop Reporting Unit Statistics

Agricultural statistics may be useful in the identification of wheat, particularly when the ratio of acreage planted in wheat differs significantly from that of other crops. A hypothetical situation in which the agriculture statistics are useful in identifying spring wheat are discussed below. This case is ideal and will probably never be encountered in actual practice. Care should also be taken to determine that the county statistics are actually representative of the segment before weighing them too heavily in the interpretation process.

The agriculture statistics for a county list the acreages for spring wheat and the major confusion crops as follows:

Crop Class	Acreage	Percent County
Spring Wheat	1 600 000	60.0
Corn	800 000	30.0
Alfalfa	150 000	1.5

The Analyst has identified three dominant "color classes" in the sample segment that have acreage proportions similar to those listed above. Sixty percent of the fields are interpreted as being mature (green signature), 30 percent are a light red signature, and 1 percent are a bright red. Based on these proportions the following associations may be made:

Crop Class	Signature	Percent of Segment
Spring Wheat	Green	60.0
Corn	Light Red	30.0
Alfalfa	Bright Red	1.0

This sample shows an ideal situation where the ratios of each crop differ significantly and where the classification conforms to these ratios. In most interpretation instances, this ideal situation will not occur; however, agricultural statistics will provide a valuable input to the interpretation process.

2.3.1.5 Signature Labeling and Boundary Delineation

Wheat and nonwheat signature identification and labeling procedures are described in this section. In addition, the procedures for the delineation of boundaries in special cases are outlined.

2.3.1.5.1 Wheat Labeling

Separate as many distinct color classes on the CIR imagery as possible and identify the signatures for wheat and/or small grains. The cluster map, if available, may be used. Select and delineate as many fields as necessary to acquire these various classes. Every signature should be represented.

The fields should be located throughout the image area, and at least 500 pixels of wheat and/or small grains should be selected. In areas where fields do not meet the minimum field size (that is, fields smaller than four pixels) groups of pixels which appear to be the same crop may be used as a single training field. These pixels should be of the same subclass.

2.3.1.5.2 Nonwheat Labeling

To provide an adequate sample for computer classification, examples of all nonwheat signatures must be selected. These include both areas that are identifiable (i.e., pasture, water, cities, etc.) and areas that cannot be identified.

The unidentified areas have a distinct signature present (not obscured) but no determination can be made as to their identity. (See appendix B for further definitions and coding.)

a. Select at least 100 pixels for *each* "Nonwheat" color class present.

b. Separate these color classes into Agriculture, Non-agricultural or Unidentifiable.

c. Place the field name, type, subclass name, and description, if applicable, on the AI temporal crop interpretation form.

2.3.1.5.3 Exclusion Are Labeling

The Analyst may find it necessary to place training fields within DO and DU exclusion areas.

a. Training fields within DO area

This option is most useful when numerous small areas of the same feature lie outside a DO area. For example, a training field placed within a large lake (DO area) would be used by the computer to properly classify numerous small ponds scattered throughout the rest of the image. This may eliminate the necessity of defining all the small ponds (see fig. 2-8). In this case, the Analyst could have submitted the entire lake as both a DO area and as a training field instead of selecting a portion of the lake. This would provide a more comprehensive training signature to properly classify the small ponds.

The coding of a training field within a DO area is the same as any training field with its own unique number.

b. Training field within DU area

This option is used when numerous small areas of the same obscuration feature lie outside a DU area. For example, a training field placed within a large cloud (DU area) would be used to properly classify numerous small clouds scattered throughout the rest of the image. This may eliminate the necessity of delineating all the small clouds (see fig. 2-9).

A training field within the DU area has its own unique category and class (see appendix B).

Caution: Obscured features may not be uniform across the scene; therefore, training fields from another obscured area may not adequately train the classifier.

2.3.1.5.4 Test Fields

For each acquisition for which training fields are selected, the Analyst will select test fields. These fields are used by the Analyst as a supplemental test of the computer classification accuracy.

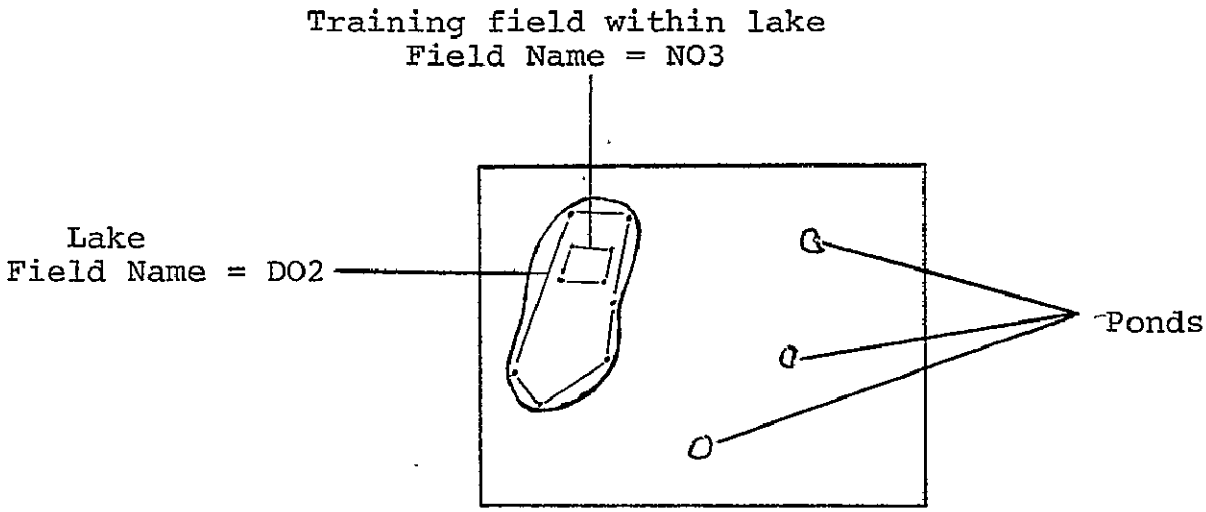


Figure 2-8. — Example of training field placed within a DO area.

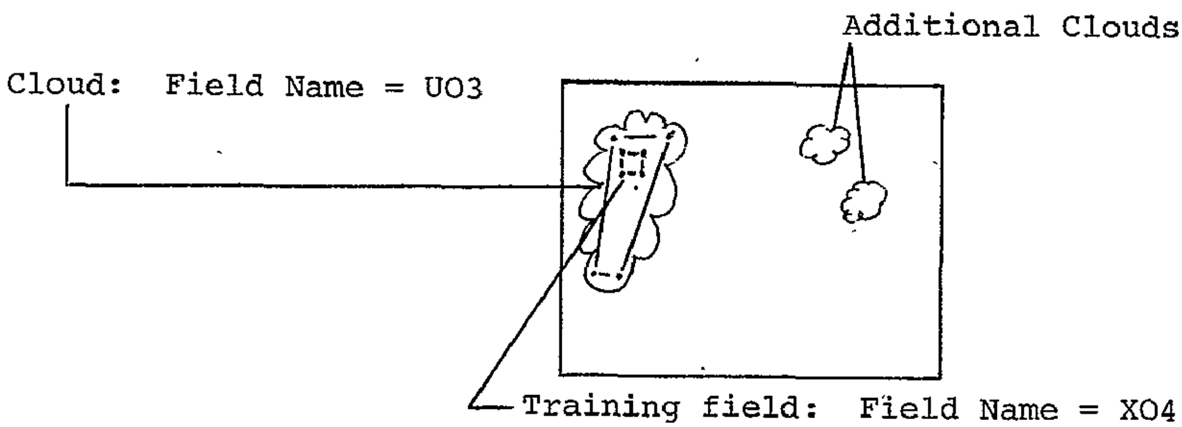


Figure 2-9. — Example of training field placed within a DU area.

If DO or DU fields do not exist on a segment, five test fields, representing as wide a range of signatures as possible, should be selected. Depending on the amount of DO or DU area in a segment, a proportionate number of test fields, not exceeding five, will be selected. Test fields will not be placed in DO or DU areas. Test fields cannot be duplicates of training fields and each test field should contain pixels belonging to only one category.

2.3.1.5.5 Boundary Delineation

To delineate boundaries, place "dots" within the field (at each corner), then connect the dots with straight lines (fig. 2-10). Previous year's imagery should be correlated with current imagery for additional information. Assign a unique sequential identifier to each delineated field (i.e., 01 to 99, 0A to 9Z), prefaced by its category designation and label on the CIR primary overlay (W01, W02, N03). Record the identifications on the AI Temporal Crop Interpretation Form. No crop identifications are made for test fields.

2.3.1.5.6 Special Cases

Fields that have visible roads, canals, or other significant features passing through them will be separated into two or more fields (fig. 2-11). This also must be done when a field is split by noise lines (data dropouts).

As special cases arise they will be worked and procedures documented.

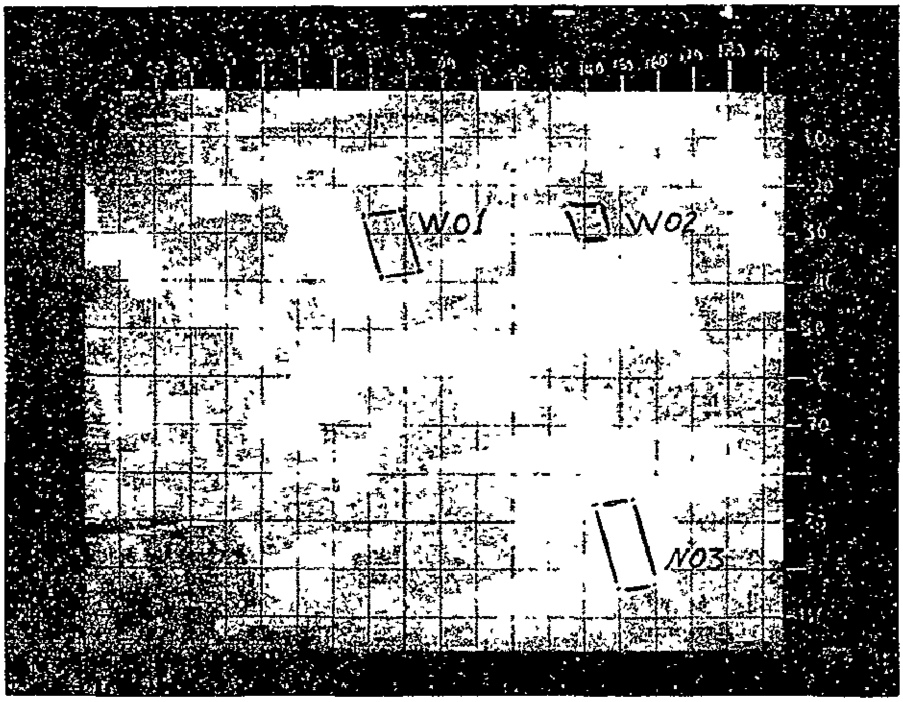


Figure 2-10. — Field boundary delineation. Annotation. W01 shows the area within a wheat field that was selected as a wheat training field. The dots with straight lines connecting them show the corner coordinates of the training field. Additional sequential identifiers prefaced by the category are shown.

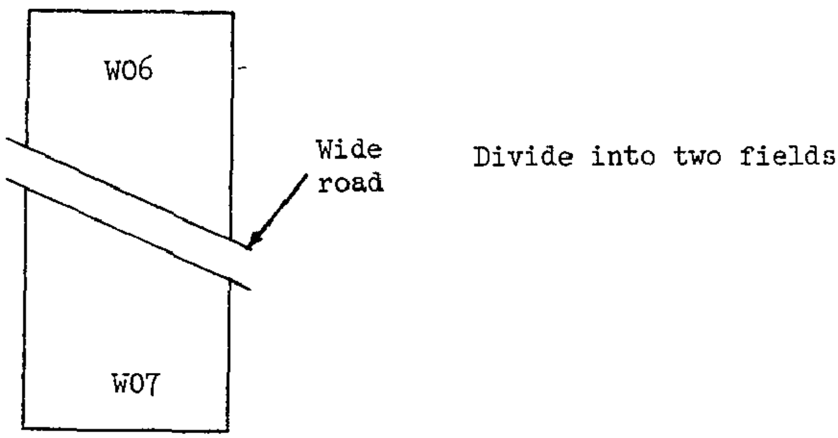


Figure 2-11. — Field divided by road.

2.3.2 MULTITEMPORAL ACQUISITION CLASSIFICATION

The procedures for this task are the same as those outlined in section 2.3.1 except for the following procedures.

2.3.2.1 Signature identification

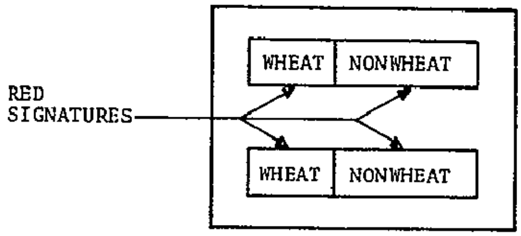
The conditions necessary for this decision are detailed in section 2.2.3. This situation usually occurs when the most recent acquisition has good wheat signatures, but they are confused with nonwheat. The Analyst will select multitemporal signatures from two, three, or possibly four acquisitions. The signatures selected for wheat and non-wheat training must be selected and placed into unique subclasses that separate wheat and non-wheat.

The procedures for selecting typical multitemporal signatures are:

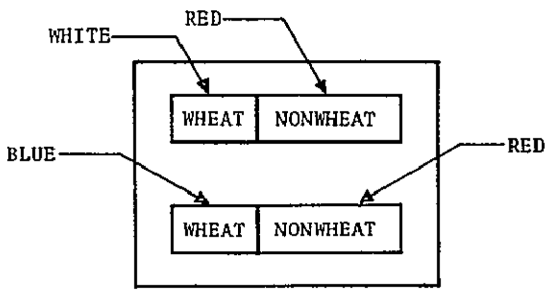
(1) Select the acquisition with the most uniform and homogeneous signatures (usually the most recent acquisition). See figure 2-12a. Wheat and nonwheat signatures are confused and the Analyst will have identified each signature through the temporal interpretation of spectral and spatial information from two or more acquisitions.

(2) Compare the confused wheat/nonwheat signatures (spatially) to all earlier acquisitions.

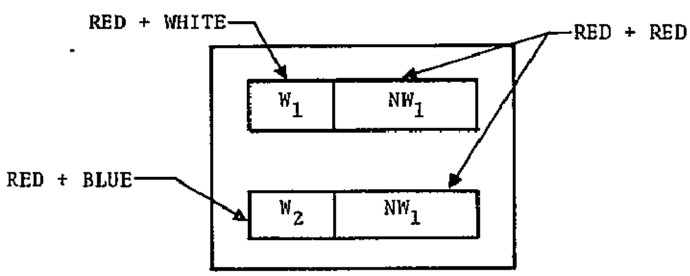
(3) When the wheat/nonwheat can be separated by the addition of spectrally different signatures from an earlier acquisition as illustrated in figure 2-12b, the multitemporal



(a) Most recent acquisition with good wheat signatures but confusion of wheat/ non-wheat.



(b) Prior acquisition with poor wheat signatures, but separation of wheat/ non-wheat.



(c) Multitemporal (8 channel) signatures from the combination of two acquisitions.

Subclass	Signatures
W_1	Red + White
W_2	Red + Blue
NW_1	Red + Red

Figure 2-12. – Selection of multitemporal (8 channel) signatures.

signatures will be delineated and placed into unique sub-classes as shown in figure 2-12c.

2.3.3 MECHANICAL PREPARATION FOR AUTOMATIC DATA PROCESSING

The line/pixel coordinates for the training fields must be recorded on punch cards and transmitted to building 30 for updating the field data base. These procedures are given below.

2.3.3.1 Corner Coordinate Recording

See appendix D for the detailed procedures.

2.3.3.2 LARS Reformatting and Editing

See appendix D for the detailed procedures.

2.3.3.3 Reformatting User ID Card for Field Update

The user ID card from the FLAP deck is replaced by a reformatted user ID produced on the Laboratory for Applications of Remote Sensing (LARS) terminal by the CAMS data clerk. An example of the reformatted data is shown below.

USER 1-3315-3635-CWH-04-C977A

- USER → Identifier Code
- 1 → Biophase
- 3315 → Acq. Date
- 3635 → Current Date
- CWH → Analyst's Initials
- 04 → DPR Code
- C977A → DPR Number

2.3.3.4 Batch Classification Request Card Generation

The user ID card from the FLAP deck is used to generate the Batch Classification Request cards automatically on the LARS terminal. Section 3.1.2, EXAMPLE BATCH INPUT DECKS, described the cards which will be generated. NOTE: Card decks for

multitemporal classifications are not produced automatically. Instructions for manual coding are given in section 3.1.1.

2.3.3.5 Job Submittal

FLOCON submits the Fields Update Deck and the Batch Classification Request Deck (with DPR's) to LPDL.

2.3.4 POST INTERPRETATION PROCEDURES

a. Place the following items in the team packet.

   (1) Imagery, ancillary summary, maps, and crop calendar.
   (2) Original Temporal Crop Interpretation Form
   (3) FLAP printout (offline) into acquisition packet for that date.

b. Remove the header information from the Product 1 and place in the segment acquisition envelope. Place the Product 1 image in a document protector along with other Product 1's for storage. When acquisitions within one day are available, only one is necessary to be placed in the document protector.

2.4 TRAINING FIELD REVISION

The function of this task is to improve the wheat estimate provided by the initial training field selection procedure. Prior to beginning this task, the Analyst will have evaluated the packet contents, updated the Analyst Status Form, verified the location of the segment, prepared the imagery for annotations and become familiar with the background data.

The decision to revise the training fields is made after the Analyst has reviewed the most recent acquisition(s) in the team packet and has determined that a significant change in the wheat estimate has occurred.

The revision procedure may occur by selecting training fields from a single acquisition or selecting multitemporal training fields.

The revision of a wheat estimate through training field revision may occur more than once and subsequent revision procedures are the same as those described in this section.

2.4.1 SINGLE ACQUISITION CLASSIFICATION

The procedures for this task are the same as those described in section 2.3.1 with the following exceptions.

2.4.1.1 Wheat Labeling and Boundary Delineation

a. Annotate any field boundary changes. The size of a training field may be decreased or enlarged and its shape may also be altered to reflect signature changes.

b. Determine if the wheat identifications, based on the first training field selection, are still correct; that is, do the signatures correspond to the progression in phenological stages?

c. If any of the original interpretations remain the same, retain these training fields by annotating the boundaries and field numbers on the primary overlay prepared for the revisions.

d. Reclassify or drop any fields that no longer appear to be wheat.

e. Select new training fields to replace those dropped or reclassified, retaining an adequate number of wheat pixels (minimum of 500). NOTE: Field names of added fields cannot be duplicates of previously used field names.

f. The positions of training fields from the first training field selection process will not be changed during the revision procedure.

g. If a field is dropped, the field name cannot be used for another field.

2.4.1.2 Nonwheat Labeling and Boundary Delineation

The procedures for labeling and delineating boundaries on nonwheat areas are the same as for wheat. See sections 2.4.1.1, 2.3.1.5.2, and 2.3.1.5.5 for procedures.

2.4.1.3 Exclusion Areas Labeling and Boundary Delineation

See sections 2.4.1.1 and 2.3.1.5.3 for procedures.

2.4.2 MULTITEMPORAL CLASSIFICATION

Training field revision using multitemporal classification can be used to update a previous single or multitemporal classification. The decision to use this procedure remains the same as for the initial training field selection procedures in section 2.2.3.

The detailed procedures for selecting revised multitemporal training fields are the same as those outlined in section 2.3.2.

2.4.3 MECHANICAL PREPARATION FOR AUTOMATIC DATA PROCESSING

All procedures are the same as those in section 2.3.3 except those noted in section 2.4.3.1.

2.4.3.1 Corner Coordinate Recording

The corner coordinate recording for the revised interpretation is dictated by the changes that occur between the initial and revised interpretation. The following table indicates when coordinates must be recorded for the revision.

Field status	Record coordinates for revised interpretation	Deletion card
No change in field data	No	No
Field category change	Yes	Yes
Field subclass change	No	No
Field coordinate change	Yes	No
Field deleted	No[a]	Yes

[a]No coordinates are recorded but a "Delete" card must be typed on the H. Dell Foster Digitizer (see appendix D).

For fields that require coordinate recording, follow the procedures outlined in appendix D.

Fields that do not require coordinate recording based on the criteria above may require recording due to misregistration. When temporal misregistration occurs, each field must be reviewed individually to determine if rerecording of coordinates is necessary. For example, if the misregistration is a one pixel shift in large fields where the "dots" were placed two pixels inside the field boundary, there would be no requirement for new coordinates. However, small fields where the "dots" were placed on the field boundary would require new coordinates to be recorded.

2.4.3.2 <u>LARS Reformatting and Editing</u>

Repeat procedures outlined in appendix D.

2.4.3.3 <u>Post Interpretation Procedures</u>

Same as in section 2.3.4.

3. DATA PROCESSING PROCEDURES

All segments will be initially processed using the batch system, except segments that are determined to be processed interactively only by the Analyst Team. Initial classification will be a subclass statistics classification to the category level, with equal *a prioris* and threshold equivalent to 1 percent. For initial runs, the Analyst will have determined whether the current acquisition alone will be classified or whether a multitemporal analysis using the current acquisition in conjunction with one or more previous acquisitions will be classified. (See section 2.2.3, MULTITEMPORAL DECISION CRITERIA for more detail.)

For rework, the full batch capabilities and/or the interactive system may be used in accordance with section 3.4.5, PROBLEM DETECTION GUIDELINES.

3.1 BATCH RUN PROCEDURES

This section describes the procedures for obtaining the "summary of summary" and the submittal of a multitemporal batch classification.

a. CAMS personnel daily submit a job to process the CAMS Report Tape Computer Output Microfiche (COM)[1] Tape which contains batch results of the previous night. A printout containing the statistics report, the classification summary, and the results of the calculations described in section 3.4.7.2 is the final product. A detailed description of the cards and forms necessary for this job is included in appendix D.

b. To perform a multitemporal classification, indicate the required acquisition dates on a note accompanying the FLAP deck. The CAMS Data Clerk can then prepare the necessary cards.

---

[1]Plans are to use the CAMS/CAS Interface Tape in the future.

3.1.1 BATCH INPUT CARD FORMATS

To access the processing options in batch mode, a set of cards is input to process control. These are the cards that are generated on the LARS terminal as described in section 2.3.3.3. There are 19 card types which are allowed. Of these, three are used in the nominal batch job. Multiple cards of most types are allowed. A detailed explanation of card types, restrictions, and usage can be found in the *LACIE User's Guide*. Knowledge of these three cards is used to debug batch classification runs which terminate abnormally.

This section will give an abbreviated description of the three common cards, examples of each card, naming conventions, and restrictions. Also three examples of complete jobs are given. The remaining cards are described in appendix E, PROCESSOR CAPABILITIES, with the processor to which each applies.

<u>Card 1</u>: NAME=XXXXX

Up to 24 characters may be input. This card must be the first card of the batch input deck.

```
                                    Date of processing or date
                                    DPR originated.

                                    Initials of originator

                                    DPR code

Example:  NAME=X-3145-4347-CRS-07-A077 ← DPR#

          Biophase                  Acquisition date 3145 placing
          X=1-4                     late digit of last year.
                                    Example:  3425 for 12/8/75
```

<u>Card 2</u>: RSEG=nnnn,nnnn,(nnnn),nnnn

This card indicates the site numbers of the sites to be processed in the manner indicated by the remaining cards in the batch deck. The parentheses indicate that the training fields for this site

would be changed to test fields, and therefore, statistics are not computed. It should be noted that the inputs are numeric site ID's, not preceded by an "I". Also, individual report fiche will be generated for each recognition segment, with the same eyeball readable name. The image merge channel correspondence report will indicate which recognition segment is being classified and all other reports on that fiche refer to that segment.

Example: RSEG=1108,1843,(1210),1560

Card 3: CSEG=nnnn(yyddd,yyddd)
CSEG=(yyddd,yyddd)

This card specifies the dates of acquisition to be used in image merge for sites listed on the RSEG card. If no site number is listed on the CSEG card, the dates apply to *all* sites on RSEG card except for those which another CSEG card is present. For example,

```
    CSEG=(75010,75102)
    CSEG=1108(75009,75102) is a valid combination,
but CSEG=(75010,75102)
    CSEG=(75009,75102) is not a valid combination.
```

The same number of dates must appear on each CSEG card.

On all batch card inputs, except Card 1, no imbedded blanks are allowed. If cards need to be continued, column 80 should be coded and the keyword repeated on the next card.

3.1.2 EXAMPLE BATCH INPUT DECKS

a. To classify two segments using four acquisition dates each:

```
NAME = 1-3145-5085-CAS-07-A999B
RSEG = 1106,1034
CSEG = 1034(74056,74128,74146,74182)
CSEG = 1106(74056,74128,74163,74182)
```

b. To classify a single segment with single acquisition date:

```
NAME = 1-3145-5119-CAS-07-A766B
RSEG = 1106
CSEG = (74182)
```

c. To classify a single segment with three acquisitions:

```
NAME = 1-3145-5130-CAS-07-A606B
RSEG = 1106
CSEG = (74056,74128,74182)
```

3.2 CAMS INTERACTIVE RUN PROCEDURES

This section describes the procedures for an interactive run to be made on the LACIE system operating on the IBM 360/75. Interactive processing is primarily used for reworking segments that previously have been processed batch. Inputs are changed based on the nature of the problem.

When it is deemed necessary by the Analyst that a segment should be processed interactively, time is scheduled through FLOCON. Details are to be designated.

The screening imagery and possible processing suggestions should be reviewed by the Analyst prior to going on the terminal. Checks are made for corrections or changes as described in section 3.4.5.

Details of the console operation are documented in the *LACIE User's Guide*. Overviews and the theory of each processor are in appendix E.

3.2.1 SOFTWARE CONSIDERATIONS

A DPR number must be obtained prior to the scheduled interactive time.

a. Sign On - To sign on the terminal, hit the "Color Release" key. The next menu to appear will ask for name, ID, room number, building number, FOSO number (811), and PA number (3783). On the first line, NAME, the first 4 characters are used to identify DELOGS; therefore, the first four of the last five characters of the 24 character DPR number should be entered here. It is mandatory to put the complete DPR number on the second line (ID) because this line is input to status and tracking. Select Pattern Recognition.

b. Image Merge - The merged image should be name Mnnnnx where nnnn is the segment number and x is the next available alphabetic for that segment number.

c. Clustering and Chaining - Output names must not duplicate names already in the system. Names should be consistent with the categories on the data base. Specifically, wheat statistics should begin with the letter W, and nonwheat should begin with N. The second character should be converted to the proper numeric to correspond to the segment location, i.e., 0-3.

d. Classification - Check Option 4, the RSEG box, on the classification menu to have the segment classified with the special processing.

e. Class Map - Type in the following color codes for color image or classification image tape.

Unclassified Color Number	0 (black)
Category Wheat	14 (white)
Category Nonwheat	8 (red)
Category Unidentifiable	6 (cyan)

f. General - If the terminals get "hung" in an application, the LACIE Data Systems Supervisor (LDSS) should be informed.

He will communicate to the user and the operator the best method for correcting the situation while retaining maximum user information and data base integrity. The DPA must bring the system down with a STOP between analysts in order to have a new DTERM tape, which contains all Production Film Converter (PFC) images mounted; SIGNOFF will not do this.

g. Products - A checklist of the possible outputs includes:

a. Hard copy

- Field definition
- Mean/standard deviation
- Feature selection report
- Cluster report menu
- Detailed clustering report
- Intercluster distance table
- Detailed classification summary
- 64-gray shade map

b. Microfiche or print

- Classification map
- Cluster character map
- Delog

c. Tapes

- Fields tape (update)
- CAMS/CAS interface tape

d. Film Products

- Color cluster map
- B/W classification map

3.2.2 INTERACTIVE POST PROCESSING INTERFACE PROCEDURES

The operator and/or Data Production Control Analyst (DPCA) should be notified and given the DPR number for each terminal in order to separate products. If a DPA generates a nonstandard tape, i.e., a signature extension tape or image tape to be saved, the

DPA should tell the operator to save it and what the tape contains. If the DPCA does not provide a label, then the DPA should take a save label to the operator.

To obtain a delog, a DPA should delog on the system via the delog menu or check the delog on the DPR under COM products. Pattern Recognition *reports only* should be checked with the times desired. Unless otherwise notified, these reports will be the only thing put on the fiche delog. If the delog is done on the system, the run should be brought down 10-15 minutes per console, depending on the length of the run, before the end of the scheduled time. If a delog is checked on the DPR then the delog will be on microfiche. The fiche can be obtained from LPDL when processing is completed.

If any character maps (print) are generated, the DPA should collect the maps along with the paper delog at the computer after the system is brought down.

At the end of the run, the DPR should be completed to correspond to the products actually generated and four copies forwarded to FLOCON.

3.3 PROCEDURES FOR SPECIAL CASES

Certain situations may exist where wheat is separable, but the fields are of a shape and size that makes it impossible to enter them into the Fields Data Base. Since these cases are unique to individual segments and cannot be generalized, a detailed approach is not discussed here. However, as these cases are worked, each successful procedure will be documented and made available for future reference.

3.4 EVALUATION PROCEDURES

This section describes the methods by which an assessment is made of the end product, which is an electronic classification (emphasis being on wheat) using remotely sensed data, by the Analyst Team.

3.4.1 PRODUCTS ESSENTIAL FOR EVALUATION

The following products from batch or interactive processing are needed for the proper evaluation of the segment:

a. Classification summary and/or 1110 printout (summary of summary)

b. Statistics - Mean and Standard Deviation Report, Detailed Clustering Report of Chaining Report (whichever is applicable)

c. Field Definition Report

d. Classification Map

e. Cluster Map (if applicable)

3.4.2 EVALUATION FLOW PROCEDURES

The following steps are to be taken by an Analyst in order to ensure proper evaluation of a segment.

a. If the Analyst decides that the current acquisition has no significant change as described in seciton 2.1.7 or has less than 5 percent wheat/small grain contained in the segment, then the CEF is completed as described in section 3.4.6.1. This completed form is passed to CAS via FLOCON.

b. If training fields for this acquisition have been submitted, classify the segment using standard classification procedures as defined in section 3.1.

c. The Analyst should determine if the classification results are satisfactory, marginal or unsatisfactory utilizing the evaluation criteria in section 3.4.3.

d. If rework is necessary, the Analyst should rework the segment using the problem detection guidelines until the segment is judged satisfactory, marginal, or unsatisfactory, or until the Analyst has reworked the segment batch once and one interactive session.

e. The Analyst should then complete the CEF with the required information as described in section 3.4.7.1.

f. The Analyst places one copy of the CEF in the Analyst's Packet along with all of the products from classification. He then returns the packet to the Data Clerk and submits three copies of the CEF to the Data Clerk for distribution to CAS and appropriate files.

3.4.3 EVALUATION CRITERIA

The Analyst should check the following points before evaluating all classifications runs. Each point will be addressed further in the rating of a segment or the Problem Detection Guidelines.

a. Individual training field accuracy at the category level.

b. Training class accuracy at the category level.

c. The percent threshold for individual training fields.

d. The statistics used for classification (i.e., the range of means and standard deviations).

e. The percent threshold for the entire segment.

f. The usage of DU.

g. The correlation between product 12 and the primary overlay.

h. The performance of the test fields.

i. The correlation of the class map and the original imagery.

j. Nonsimilarity of signatures of high standard deviations within a single subclass.

k. Test fields located in DO or DU areas.

1. Misclassification of any obvious nonwheat area (as determined by spatial information).

Having checked these points, the Analyst will have assimilated the necessary information for the assignment of a rating to the classification.

3.4.3.1 Satisfactory Rating

The criteria that a classified segment must fulfill to obtain a rating of satisfactory are given below.[1]

a. The classification map should show definite correlation to the original imagery (i.e., those areas felt to be wheat on the imagery are classified as wheat and those areas felt to be nonwheat on the imagery are classified as nonwheat.)

In addition, at least four of the following criteria must be met:

b. All training classes should have 90 percent accuracy by category.

c. Each training field should be classified at least 80 percent correct by category.

d. The percent thresholded for the entire segment should not exceed 5 percent and should consist mainly of nonagricultural areas.

e. The percent thresholded for any training field should not exceed 5 percent.

f. All test fields should have at least 80 percent of their pixels classified into the same category; if a test field is not homogeneous, then its classification should be representative of the imagery and may be accepted. (See paragraph 2.3.1.5.4 for test field selection procedure.)

[1] The estimates obtained from a manual count will be given a satisfactory rating.

3.4.3.2 Marginal Rating

The marginal rating is used when a segment does not meet the criteria for a satisfactory rating, but all the criteria below are satisfied.

a. The classification map is acceptable if it shows a definite correlation to the original imagery; however, there can be scattered pixels that are not felt to substantially change the appearance of the entire classification map relative to the original imagery.

b. All training classes should have 70 percent accuracy by category.

c. The percent thresholded for an entire segment can exceed 5 percent only if the majority of thresholded pixels are on the boundary of DO areas, and/or if the signatures which are thresholded are nonwheat.

d. The training fields that have accuracies less than 80 percent are insignificant if:

   (1) The signatures represented by these training fields represent a significantly small portion of the original imagery (less than 500-1000 pixels).

   (2) The error can be adjusted or noted on the CEF.

   (3) The subclass statistics are still representative of the other fields in that subclass, regardless of the inclusion of the poorly classified fields.

e. Each test field should have at least 70 percent of its pixels classified into the same category or be indicative of the original imagery.

3.4.3.3 Unsatisfactory Rating

If any segment does not satisfy the criteria above for marginal or satisfactory rating, an unsatisfactory rating is given to that classification.

In addition to the above criteria, the Analyst should also consider acquisition dates, data quality, previous acquisitions, and separability of classes on imagery when making an evaluation. This evaluation rating must then be transformed into one of the codes as explained in section 3.4.6.2 and recorded in the appropriate box on the CEF (section 3.4.6.1).

3.4.4 CRITERIA FOR REWORK

Segments are eligible for rework when the Analyst feels that after classification confusion still exists between wheat and nonwheat on the classification map and that the current classification is unsatisfactory. The Analyst may consult the Regional Analyst and/or the Country Specialist, if he feels that the evaluation of the current classification indicates that the wheat estimate for the segment can be improved substantially by use of the PROBLEM DETECTION GUIDELINES, section 3.4.5. Segments should be reworked no more than once using batch processing and once using interactive processing.

3.4.5 PROBLEM DETECTION GUIDELINES

3.4.5.1 Classification Map

The classification map is the most immediate indicator of the success of the classification attempt. The classification map should display a field pattern which reasonably corresponds to the segment imagery and the field overlay. Wheat may have the same signature on one phase as nonwheat; and it then becomes necessary to look at imagery from a previous phase to distinguish the wheat areas, or conversely, the nonwheat areas.

Trouble areas (such as poor training or test field accuracies, high thresholding, and bad data) are usually visually apparent on the classification map and alert the Analyst to possible areas of concentration (listed below according to the nature of the problem) for detailed analysis. Sometimes streams, forests, and other areas (if not DO'ed) are often confused with wheat in spite of training fields; and thus, if this confusion is detected by the Analyst on the classification map, the map can serve as a guide to the placement of additional DO fields for rework or multitemporal analysis.

A note of caution to the Analyst is that training field accuracies which satisfy the requirements of the evaluation procedures do not independently reflect a good classification. The major portion of the evaluation is based on the classification of the entire segment, as reflected in the classification map.

3.4.5.2 Thresholding

The pattern of the thresholding, as well as the amount of thresholding on the classification map, can serve as a tool for possible problems within the segment and probable solutions.

Uniform thresholding around field boundaries usually points to inaccurate coordinate placement by the Analyst due to misregistration, especially if multitemporal processing is being attempted or if field coordinates extracted from one acquisition are used on a different acquisition. When thresholding of entire fields or large unified areas occurs, the imagery should be checked for clouds, cloud shadows, or haze, for which no training or DU fields exist. The Analyst should be especially alert that these phenomena are properly treated. Once the Analyst is assured that unidentifiable fields have been handled, but the above problem still persists, it can be an indication of insufficient or inadequate training classes or subclasses. This can be checked by

reviewing the imagery. If it can be determined upon review of the imagery that only nonwheat signatures are missing, then the wheat estimate might be accepted without reprocessing, depending on other evaluation criteria. If potential wheat signatures are missing, then additional training fields should be added. Inadequate training classes should be refined. Thresholding in excess of 10 percent for the entire segment normally is caused for rejection of a classification attempt.

3.4.5.3 <u>Training Field Accuracy</u>

All individual training fields should have at least 80 percent of the pixels allotted to the proper category, and individual classes should have at least a 90 percent classification accuracy to the category level. If this condition is not satisfied, the training statistics may be nonrepresentative or mislabeled, or coordinate inaccuracies may be involved. If the classification mistakes are restricted to a small number of fields, it may indicate that the subclasses used were not representative of these fields. The Analyst may try to form new subclasses through changes in clustering or chaining in order to obtain more representative subclasses.

The imagery, along with the temporal form, should be checked also for dissimilarity of fields grouped into the same subclass. If this condition exists, then the statistics might not be representative of any of the fields. The subclasses should be refined.

3.4.5.4 Test Fields

Each of the test fields is expected to be homogeneous and should classify with at least 80 percent of its pixels in one category. If all signatures have representative statistics, then each test field ought to be very similar to a training field and should be classified into the same category. Classification of the test fields reflects the training of the classifier.

If test fields are located in DO or DU areas, it should be noted; they are expected to have a considerable amount of threshold pixels. If they are not located in these areas and have poor classification accuracies, then this is an indication that training fields need to be added, refined, or redefined (regardless of training field accuracies). If the field was not homogeneous, the classification map should reflect the imagery.

3.4.5.5 Imagery and Field Overlay

The training and test field overlay should be compared to the imagery. If the field overlay does not fit the imagery, then there is a high possibility of coordinate mistakes. Such mistakes can be caused by mistakes in Analyst-produced work or misregistration. To check the registration, the Analyst should compare imagery which has been previously processed with the field overlay. If fields on the overlay fit the image for one pass of data but not another, there is a high probability of misregistration or change of coordinates. Any mistakes made by the Analyst should be corrected before reprocessing.

All imagery should be checked for streaking or other signs of bad data, and such information should be noted in the proper files. If bad data is restricted to certain channels, classification should be made without using those channels with the bad data.

3.4.5.6 Statistics

Class and subclass means for agricultural areas should be checked for values greater than 70 or less than 5. Such values indicate the possibility of cloud or cloud shadows in the class or subclass, or possible bad data. However, it should be noted that other fields such as sand dunes, or harvested crops sometimes have high means. These normally should be identified as such by the Analyst.

If class or subclass standard deviations have values greater than 10, it is likely that a single class or sublcass contains either bad data or diversity within the class or subclass. To remedy this situation, the Analyst should first double check for bad data, or possible overlay of a field into an adjacent area and then attempt to rework the statistics by redefining subclasses or clustering to perhaps split the class or subclass further.

3.4.5.7 Clustering Report and Map

Clusters which represent only a small portion (less than 10 percent) of the training pixels for the class from which these clusters were derived can bias classification if given the same weight as large clusters. In order to prevent such mistakes, these clusters should be checked on the cluster map. If the cluster is spread about and consists of mostly boundary pixels, the Analyst should check the training field boundaries. If there is doubt concerning the field identity of such pixels, the small cluster can be deleted or the Analyst may consider redefinition of the field. If a small cluster takes the form of a small number of individual fields, the Analyst should note any discrepancies in the class match on the CEF.

If a small cluster has more than 10 percent of the entire segment classified into it, it should be noted in the evaluation.

If the Analyst obtains a large cluster which is closest to another class, he should check the cluster map to determine if the cluster is made up of a group of fields. This should also be noted in the evaluation.

If whole-image clustering has been performed, the cluster map may be referenced for location of common signatures and compared to the subclass definition of training fields. This is especially helpful in multitemporal situations for detecting changing signatures.

3.4.6 RESULTS ARCHIVAL AND TRANSMITTAL

This section displays and explains the inputs on the form that is used to record the results of a classification before it is passed to the CAS; the data is recorded on the packet Data Acquisition/Evaluation Record before the packet is submitted to the QA team.

3.4.6.1 CEF

This form, figure 3-1, is the primary means for recording and archiving the disposition of segments. A CEF is completed each time a classification is performed on a segment (acquisition) and a wheat estimation is to be passed to the CAS. If a new wheat estimate is not necessary or possible, see section 3.4.6.3.

The following is a description of the information recorded in reference to numbered blocks in figure 3-1:

1. Segment number being classified
2. Predesignated biophase
3. Julian date of acquisition being processed
4. Last digit of the year the acquisition was assigned
5. Code from CAMS Evaluation Codes (section 3.4.6.2)
6. Date the Analyst completes evaluation for the specified acquisition.
7. Names of all team members involved in interpretation and/or evaluation. Both team members should sign the form.
8. DPR number

9. Physical location of the segment (for example, Morton, Kansas; Alberta, Canada)
10. Analyst's interpretation of biophase based on actual imagery
11. This box is checked when wheat has not emerged; it corresponds to code 07 of the CAMS Evaluation Codes.

CAMS EVALUATION FORM

SEG# 1 ☐☐☐☐ 2 ☐ 3 ☐☐☐ 4 ☐

RESULTS 5 ______ 6 __/__/__

ANALYST(S) 7 ______

DPR # 8 ______

LOCATION 9 ______

AI BIOPHASE 10 ______

NW IDENTIFICATION: 15

11 ☐ PRE-EMERGENCE

12 ☐ CLOUDS/HAZE/SNOW

13 ☐ NO SIGNIFICANT CHANGE

14 ☐ LESS THAN 5% WHEAT/ SMALL GRAINS ______ PIXELS

ACQUISITION DATES

16 ______

______

______

______

17 ☐ FEATURE SELECTION ______

18 ☐ STATISTICS MANIPULATION ______

ACCURACY

			TEST					TRAINING		
			%	CAT	%TH		%	%TH	NAME
	19	P 25	26	27	28		29	30	31
W	19	P 25	26	27	28	W	29	30	31
W	19	P 25	26	27	28	W	29	30	31
W	19	P 25	26	27	28	W	29	30	31
W	19	P 25	26	27	28	W	29	30	31
N	20					N	32	33	
DO	21								
THRSH	22								
X*X	23								
DU	24								

Evaluation Form Revised 12/12/75

Figure 3-1. – CAMS Evaluation Form (CEF).

12. The box is checked when the segment is not processed due to clouds, haze, etc.; corresponds to code 01 of the CAMS Evaluation Codes.

13. This box is checked when there is no significant change in the wheat estimate from a previous acquisition; corresponds to codes 12, 22, and 32 of the CAMS Evaluation Codes.

14. This box is checked when the Analyst feels there is less than 5 percent wheat/small grain on the imagery. The number of wheat/small grain pixels is estimated and converted to a wheat estimate on lines 19.

15. All nonwheat which the Analyst has identified on the Temporal Crop Interpretation Form should be noted (for example: pasture, fallow, barley, corn, alfalfa, lake, salt flat).

16. Acquisition dates for all passes being classified.

17. If a subset of channels are used for classification, the box is checked and channels used are noted.

18. If any changes are made to the fields or subclasses which are on the data base, sufficient information should be noted in order for the run to be duplicated.

19.
$$W{*}XX\% = \frac{W_c + \dfrac{W_c\ (X_c + \text{DU pixels})}{22932 - \text{DO pixels} - \text{DU pixels} - X_c}}{22932} \times 100$$

W_c = number of pixels classified into class W*XX.
X_c = number of pixels classified into category X.

20. $NW\% = 100 - \sum_{i=1}^{n}$ percentage of wheat in WXXX class

where n = number of wheat classes

21. $$DO\% = \frac{\#\text{DO pixels} \times 100}{22932}$$

22. $$\%\ \text{Threshold} = \frac{\Sigma\#\text{pixels thresholded at category level} \times 100}{22932 - \text{DO pixels} - \text{DU pixels}}$$

23. $X{*}X\% = \dfrac{\#\text{pixels in category X} \times 100}{22932}$

24. $DU\% = \dfrac{\#\text{ DU pixels} \times 100}{22932}$

25. Test field number

26. $\text{Test } \% = \dfrac{\#\text{ of "correct" pixels x } 100}{\#\text{ of pixels in test field}}$

27. The category which receives the most pixels in that field is assumed to be the correct category.

28. $\%\text{ threshold} = \dfrac{\Sigma\#\text{ of pixels thresholded at category level for that field x } 100}{\#\text{ of pixels in that field}}$

29. Wheat training field accuracy is computed for each wheat class to the class level

$$\frac{\sum_{i=1}^{n} (\#\text{ of pixels of field i classified into correct class}) \times 100}{\sum_{i=1}^{n} (\#\text{ of pixels in training field i})}$$

where n = number of training fields in the class

30. $\%\text{ threshold} = \dfrac{\sum_{i=1}^{n} (\#\text{ of pixels threshold from field i at category level} \times 100}{\sum_{i=1}^{n} (\#\text{ of pixels in training field i})}$

31. Wheat class name — If cluster or chain statistics were used for classification, the class names must be converted to codes CAS can understand (i.e., a code from the standard codes).

32. Nonwheat is calculated to the category level

$$\frac{\sum_{i=1}^{n} (\text{\# of pixels of field i classified into category N}) \times 100}{\sum_{i=1}^{n} (\text{\# of pixels in training field i})}$$

where n = number of fields in category N

33. Same as #30 except that n is the number of training fields in category N.

3.4.6.2 CAMS Evaluation Codes

The codes are used to transform the satisfactory, marginal, and unsatisfactory ratings and nonprocessed acquisitions to a numerical representation that is interpretable for CAS proceedings. This code is required in line 5 on figure 3-1.

The following codes were developed in response to concerns by CAMS, CAS, and Project Management for a more detailed set of codes that would better cover the various criteria used within CAMS to evaluate and process Landsat data.

These codes were also developed to meet certain CAS requirements for evaluating the CAMS data used for CAS area aggregations and for ease of machine processing of data.

Two sets of criteria were used in developing these codes:

a. Method of tracking and retrieving CAMS evaluation data by segment.

b. Method of accounting for data not processed because of technical problems, ground problems, etc.

A detailed explanation of the codes follows.

Code 01: Not machine processed – clouds, haze, snow, etc.

Means that the acquisition cannot be processed through the system because clouds, haze, etc., make interpretation and analysis impossible.

Code 02: Not machine processed – confusion crops or other interpretation problems.

This code should be used when an acquisition cannot be processed because of interpretation difficulties, especially when confusion between wheat/small grains and other crops is such that a wheat/small grains estimate cannot be determined.

Code 03: Bad data – due to technical problems, not reordered.

In this case the acquisition cannot be processed due to technical problems arising from bad photo processing, excessive scan line drop, etc. If the acquisition is not reordered, this code is used.

NOTE: If the acquisition is reordered, no code should be listed and no data sheet should be passed to CAS since this segment will again be sent through the CAMS for evaluation. Any acquisition that cannot be processed due to clouds, haze, snow, etc., or technical problems should be passed as 01 or 03.

Code 05: Not machine processed – dormancy.

This code is used in cases where recognition is a problem because the crop is in a state of dormancy.

Code 07: Not machine processed – preemergence.

This code is used in cases where the acquisition is prior to the criteria established for wheat recognition (to be designated).

Code 09: Not machine processed — multiple acquisitions.

Code 09 was developed to eliminate the problem of processing multiple acquisitions of a segment at the same time. When multiple acquisitions are available at one time, only one segment, which is determined to be the better acquisition for interpretation and analysis by the CAMS Analyst, is selected for complete processing. Other acquisitions are listed as Code 04 to indicate that they had been reviewed but were not processed through the system. This code enables CAS to review the segment results and account for each acquisition and how it was evaluated by CAMS.

Code 10: Unsatisfactory — unsatisfactory results for segment.

This code is to be used for any acquisition that has been processed through the system and based on the CAMS evaluation procedures is designated unsatisfactory.

Code 12: Unsatisfactory — no significant change.

This code is used when the new acquisition is evaluated to have no significant change from the previous unsatisfactory evaluation for the segment.

Code 14: Unsatisfactory — rework, reevaluated segment.

This code is used when a segment, which was previously passed to CAS, is reworked.

Code 18: Unsatisfactory — machine processed multitemporal analysis.

This code is used when more than one acquisition date is used to produce a proportion estimate.

NOTE: All acquisition dates used in processing should be listed on the CEF.

Code 20: Marginal — marginal results for acquisition.

This code is to be used for any acquisition that has been processed through the system and based on the CAMS evaluation procedures is designated marginal.

Code 22: Marginal — no significant change.

This code is to be used when the new acquisition is evaluated to have no significant change from the previous marginal evaluation for the segment.

Code 24: Marginal — rework, reevaluated acquisition.

This code is used when an acquisition, which was previously passed to CAS, is reworked.

Code 28: Marginal — machine processed multitemporal analysis.

This code is used when more than one acquisition date is used to produce a proportion estimate.

NOTE: All acquisition dates used in processing should be listed on the CEF.

Code 30: Satisfactory — satisfactory results for acquisition.

This code is to be used for any acquisition that has been processed through the system and based on the CAMS evaluation procedures is designated satisfactory.

Code 32: Satisfactory — no significant change.

This code is to be used when the new acquisition is evaluated to have no significant change from the previous satisfactory evaluation for the segment.

Code 34: Satisfactory — rework, reevaluated acquisition.

This code is used when an acquisition, which was previously passed to CAS, is reworked.

Code 36: Satisfactory — less than 5 percent manually (hand) counted.

This code is to be used for any acquisition in which the proportion estimate is manually counted rather than machine processed.

NOTE: This category would not be used if the acquisition was a rework acquisition. Code 34 should be used.

Code 38: Satisfactory — machine processed multitemporal analysis.

This code is used when more than one acquisition date is used to produce a proportion estimate.

Code 40: Segment is totally nonagricultural.

This code is used when the acquisition is evaluated as having no agriculture at all; i.e., no discernible field patterns.

NOTE: Acquisitions in an agricultural area that have a 0-percent proportion estimate are to be designated Code 36.

3.4.6.3 Data Acquisition/Evaluation Record

This record is attached behind the flap on each segment packet and will be completed for every segment acquisition. A reproduction of this form is shown at Figure 3-2. This form derives particular importance in that it will contain evaluation codes on those segments for which a new wheat estimate is not made to CAS via the CEF.

When an acquisition is received in CAMS for interpretation FLOCON will write the acquisition date on the form prior to distributing the packet. After the analyst completes his evaluation, he will fill in the evaluation code and completion date blocks; and write his name in the block titled "Initials". The packet will then be forwarded to the Quality Assurance team who will verify that the information is correct. When FLOCON receives the packet from the QA team they will extract the necessary information for the daily report to be passed to CAS.

Segment number

FLOCON	Aquisition date	Eval. code	Completion date	Initials	FLOCON	Acquisition date	Eval. code	Completion date	Initials	FLOCON	Acquisition date	Eval. code	Completion date	Initials

ASD 395

Figure 3-2.- Data acquisition/evaluation form.

ORIGINAL PAGE IS OF POOR QUALITY

PAGE: 3-26 REV. 1

APPROVED: [illegible]

DATE: 7-22-76

APPENDIX A

IMAGERY QUALITY CONTROL

Imagery quality is monitored for every segment acquisition in addition to a weekly review of a standard segment image generated by the PFC and processed by the Photographic Technology Division (PTD). Sample Segment 1065 (Haskell County, Kansas) is the current reference standard used in the weekly review. The following parameters must be checked for every segment acquisition, especially the ones chosen for training field selection. This task will be performed during the IMAGERY REVIEW PROCEDURE, section 2.1.

A.1 HISTOGRAM PROBLEMS

Correct gain and bias values of the scaling factors determine the color and tonal range of the composite Products 1 and 2 images. If any unusual coloration is apparent, such as purple or black imagery, this malfunction should be noted. Occasionally, color problems can be confirmed by checking the numerical values located near the bottom of the header information block. The values are identified by the title, B, S, 2ND BIAS, which refer to the bias, scale factor, and 2nd bias settings of the MSS channels used to generate the composites. These values vary considerably due to spectral reflectivity and scene brightness, and there is no "correct" range of values to check against. However, a previous color problem had the second or middle figure, referring to the scale factor, pegged at 42.6. This was a good recognition characteristic for this particular malfunction.

A.2 PIXEL DEFINITION

Pixel definition is the second parameter to be checked and each pixel should be distinct. Any defocussing or merging of dissimilar signatures should be noted as a pixel definition problem. Partial pixels indicate a problem.

A.3 ABUTMENT PROBLEMS

Abutment problems are identified by noncontiguous scan lines which appear as fine gaps or overprints in the image. This problem may occur repeatedly or only once in a sample segment.

A.4 DATA DROPOUTS AND/OR SKIPS

Data dropouts and/or skips appear as intermittent pixel losses in the scan line direction. Sometimes an entire scan line will be missing from one of the image products. This will be rejected. If the problem is original data, the segment will be worked.

A.5 GRID PROBLEMS

Grid problems manifest themselves in the displacement of the 10 x 10 pixel grid superimposed on the image. This displacement results in training field coordinate errors and should be rectified before working.

A.6 TOTAL NUMBER OF PIXELS AND SCAN LINES

The total number of pixels and scan lines, 196 x 117 respectively, must be checked when annotating the pixel/scan line numbers on the imagery.

A.7 PHOTO PROCESSING AND FILM PROBLEMS

Photo processing and film problems in the form of chemical stains, emulsion pinholes, film fog, etc., shall be rejected and reordered via a DR.

A.8 VERIFICATION OF TEMPORAL REGISTRATION

Imagery acquisitions received subsequent to the reference segment will be registered by comparing identifiable features in both images. The features from the reference image are to be

traced onto an overlay and then registered to the temporal image being checked. If all points or features match to within a pixel between the two images, the registration is considered good. When one or more points are shifted in either the pixel or line direction by more than one pixel, the distance is recorded and transmitted to DAPTS via a DR form (figs. A-1 and A-2).

A.9 CLOUDS, SHADOWS, AND HAZE EVALUATION CRITERIA

The evaluation criteria for determining whether an acquisition is workable when it contains clouds, shadows, or haze are as follows:

a. Determine if previous imagery exists for segment.

b. If so, does the imagery provide sufficient data to determine if agriculture may possibly exist in the obscured area.

c. If the obscured area is a nonagricultural area such as rangeland, it will be DO and the training field selection procedure will begin.

d. Should the obscured area contain agricultural patterns, the Analyst will assume the proportion of wheat in the obscured area is equal to the proportion of wheat in the clear area of the image.

e. The last criterion to be considered is the amount of obscured wheat potential area. If the obscured area is equal to or less than 40 percent, then the acquisition is considered workable. If not, reject the acquisition.

The proportion of wheat for an acquisition containing obscured agricultural and nonagricultural areas is shown below.

DISCREPANCY REPORT/PROGRAM CHANGE AUTHORIZATION			REPORT NO. 004792
1. ORIGINATOR NAME/ORG/PH NO. Richard Hinkle, CAMS	2. REPORTED IN REAL TIME ☐ YES ☐ NO	REPORTED TO: NAME/ORG/PH NO.	
3. TYPE OF PROBLEM ☐ PCA ☐ HARDWARE ☐ SOFTWARE ☐ PROCEDURES ☒ OTHER Misregistration of PFC images		4. TYPE OF ACTIVITY/RUN NO/CONFIG NO. LACIE Operations	
5. RESPONSIBLE AREA ☐ CDC ☐ IBM ☐ LEC ☐ NASA ☐ SISO ☐ UNIVAC ☐ OTHER			
6. DATE/TIME PROBLEM OCCURRED 08/05/75	7. LOCATION BLDG/ROOM Bldg. 17	8. PROJECT/SUBPROJECT LACIE	9. COMP SYSTEM/EQPT SYSTEM
10. SOFTWARE SYSTEM/VERSION	11. SOFTWARE SUBSYSTEM/VERSION	12. OTHER SYSTEMS/VERSION	13. PROBLEM IMPACT ☐ CRITICAL ☐ MAJOR ☐ MINOR
14. CORE DUMP REEL NO.	15. RESTART TAPE REEL NO.	16. LOG TAPE REEL NO.	17. ATTACHED DOCUMENTATION Polaroids of current and misregistered images

18. PROBLEM DESCRIPTION

1. PROBLEM: Registration of Segment 1623-30. When compared to the first acquisition, this segment was found to be misregistered by:

upper left corner	Up 2 pixels	Left 10 lines
upper right corner	Up 2 pixels	Left 10 lines
lower right corner	Up 2 pixels	Left 10 lines
lower left corner	Up 2 pixels	Left 10 lines
center	Up 2 pixels	Left 10 lines

2. ADVERSE EFFECTS: Prevents multitemporal classification.
3. RECOMMENDATIONS: Re-registration by Goddard or registered imagery within the same biophase.

FOR MANAGEMENT USE ONLY

19. DR CLASSIFICATION ☐ CRITICAL ☐ MAJOR ☐ MINOR	20. TESTED WITH: VERSION ____ UPDATE TAPE NO. ____	21. DEVELOPMENT PLAN REF. SUBSYSTEM ____ ITEM ____	22. SPECIAL COORD REQUIRED ☐ YES ☐ NO EXPLAIN BELOW
23. ASSIGNED TO: NAME/DATE	24. APPROVED BY: NAME/DATE	25. APPROVED BY: NAME/DATE	26. MODS SUBMITTED FOR: MODULE NAME/CSECT NAME
27. PROBLEM DIAGNOSIS/ACTION TAKEN			
28. RESPONSIBLE AREA OR COORDINATOR	29. CLOSED BY NAME/ORG/PH NO		DATE

Figure A-1. — Example of a discrepancy report.

REPORT OF MISREGISTRATION GREATER THAN ONE LINE/PIXEL

Segment __________ Bio _____ to Bio _____.

Upper left

_____ pixel _____

_____ line _____

Upper right

_____ pixel _____

_____ line _____

Center

_____ pixel _____

_____ line _____

Lower left

_____ pixel _____

_____ line _____

Lower right

_____ pixel _____

_____ line _____

Line up the features and record how far off the reference grid (first date) is in relation to the new date of imagery. For example, the features (2 roads, a lake, a river, and 2 large fields) of the reference imagery is overlaid with the later date imagery and the reference grid is found to be 3 pixels to the right and 2 pixels up in the upper right.

Thus: Upper right

3 pixel R

2 line up

Figure A-2. — Example of form used to indicate misregistration.

$$P_w = \frac{W_c + \frac{W_c\ (DU + X)}{22{,}932 - DO - DU - X}}{22{,}932}$$

where, P_w = Proportion of wheat for the acquisition

W_c = Total wheat pixels classified

DU = Designated Unidentifiable

DO = Designated Other

22,932 = Total pixels within a LACIE sample segment

X = Total pixels classified into category X.

The above six parameters must be checked in addition to any unusual problems which may degrade the image or preclude its interpretation. If the degradation is minor, the analyst should flag them to the DAPTS Manager. When the discrepancy precludes interpretation, a DR must be submitted to the Operations Control Center (OCC). The DR must include the Landsat scene number, tape number, and acquisition date along with a brief description and a polaroid to document the malfunction. A sample DR of a discrepancy which precludes interpretation by the Analyst is depicted in figure A-3.

DISCREPANCY REPORT/PROGRAM CHANGE AUTHORIZATION			REPORT NO. 004791
1. ORIGINATOR NAME/ORG/PH NO. A. Arro/CAMS/x4761	2. REPORTED IN REAL TIME [X] YES [ ] NO	REPORTED TO: NAME/ORG/PH NO.	
3. TYPE OF PROBLEM [ ] PCA [ ] HARDWARE [ ] SOFTWARE [ ] PROCEDURES [X] OTHER PFC Image Quality		4. TYPE OF ACTIVITY/RUN NO/CONFIG NO. LACIE Operations	
5. RESPONSIBLE AREA [ ] CDC [ ] IBM [ ] LEC [ ] NASA [ ] SISO [ ] UNIVAC [ ] OTHER			
6. DATE/TIME PROBLEM OCCURRED 12/01/75	7. LOCATION BLDG/ROOM Bldg. 17/Room 235	8. PROJECT/SUBPROJECT LACIE	9. COMP SYSTEM/EQPT SYSTEM
10. SOFTWARE SYSTEM/VERSION	11. SOFTWARE SUBSYSTEM/VERSION	12. OTHER SYSTEMS/VERSION	13. PROBLEM IMPACT [ ] CRITICAL [ ] MAJOR [ ] MINOR
14. CORE DUMP REEL NO.	15. RESTART TAPE REEL NO.	16. LOG TAPE REEL NO.	17. ATTACHED DOCUMENTATION Polaroid of imagery

18. PROBLEM DESCRIPTION

1. PROBLEM: LACIE Sample Segment 7129-10-295-5 generated from Tape ID: A53012 has poor image quality due to incorrect scaling and bias values.

2. ADVERSE EFFECTS: Imagery products 1, 2, and 4 are unuseable and precludes interpretation.

3. RECOMMENDATION: Regenerate the imagery using the correct scaling and bias values.

FOR MANAGEMENT USE ONLY

19. DR CLASSIFICATION [ ] CRITICAL [ ] MAJOR [ ] MINOR	20. TESTED WITH: VERSION ____ UPDATE TAPE NO. ____	21. DEVELOPMENT PLAN REF. SUBSYSTEM ____ ITEM ____	22. SPECIAL COORD REQUIRED [ ] YES [ ] NO EXPLAIN BELOW
23. ASSIGNED TO: NAME/DATE	24. APPROVED BY: NAME/DATE	25. APPROVED BY: NAME/DATE	26. MODS SUBMITTED FOR: MODULE NAME/CSECT NAME

27. PROBLEM DIAGNOSIS/ACTION TAKEN

28. RESPONSIBLE AREA OR COORDINATOR	29. CLOSED BY NAME/ORG/PH NO	DATE

Figure A-3.— Sample DR depicting a discrepancy which precludes interpretation.

APPENDIX B

CODING STANDARDS

B.1 LACIE FIELD DATA CODING

Two LACIE coding schemes are presented in figure B-1. The Analyst Field Codes are entered on cards via the H. Dell Foster Digitizer using information already recorded on the AI Temporal Crop Interpretation Form. A program for the LARS terminal converts these codes into the corresponding LACIE Field Codes. Figure B-1 and table B-1 should be referenced while reading the following descriptions. Part A of figure B-1 is the verbal description of the LACIE categories and classes. The DO (code D) and DU (code U) are not LACIE training categories, but are classifier exclusion codes.

B.1.1 ANALYST FIELD CODES

Part B of figure B-1 shows the Analyst Field Codes used by the analyst to initially code the fields on the primary overlay and the Temporal Crop Interpretation Form. These codes are used to simplify the Analyst's task of coding fields data. These codes are subsequently converted to LACIE Field Codes automatically.

B.1.1.1 Field Name

The first designation in Part B is the Analyst Category Code and the two digit alphanumeric that follows is the field number. The combination of the Analyst category and the field number constitutes an Analyst Field Name. The last two characters of

the field name are unique two digit numerics (i.e., 01 to 99). In the event that field numbers above 99 are required, the following alphanumeric system will be used:

Field #100 Use 0A

Field #101 Use 1A

PART A		PART B				PART C						
		ANALYST FIELD CODES				LACIE FIELD CODES						
CATEGORY	CLASS	FIELD NAME		SUBCLASS		FIELD NAME		TYPE	CAT.	SUBCLASS		
	Winter Wheat			W							WW	
	Spring Wheat			S							SW	
Wheat	Small Grain	W		K		W			W		GR	
	Winter Small Grain			B							WG	
	Spring Small Grain		01-99	D	01-50		01-99	T			SG	01-50
	Agriculture			A							AA	
Nonwheat	Nonagriculture	N	0A-9Z	N		N	0A-9Z		N	*	NN	
Unidentifiable		X		X		X			X		XX	
Test		P				P		P	P		TS	00
Designated Other		D				D		D				
Designated Unidentifiable		U				U		U				

Figure B-1. — LACIE coding schemes.

TABLE B-1. — EXAMPLES OF LACIE FIELD CODING

Analyst Code	LARS Converted Code	Description
W01 W1	W01 T W*WW01	Winter Wheat
W04 K1	W04 T W*GR01	Small Grain
N06 A1	N06 T N*AA01	Corn
N07 A1	N07 T N*AA01	Corn, same signature as No. 6
X08 X1	X08 T X*XX01	Unidentifiable
P10	P10 P P*TS00	Test Field
D11	D11 D	DO
U12	U12 U	DU

Field #102 Use 2A
Field #109 Use 9A
Field #110 Use 0B
Field #119 Use 9B
Field #120 Use 0C
etc.

The field category portion of the field name may change from one acquisition to another, but the last two characters must remain constant as long as the field is maintained. These numbers may not be used for any other location on the segment under any circumstances.

B.1.1.2 Subclass

The analyst subclass consists of an alphabetic and a two digit numeric code. This designates the subclass of each major category, such as wheat and nonwheat. Each class (i.e., winter wheat, spring wheat, small grain) has the potential of 50 subclasses. Each unique signature must have a unique subclass, even though the signatures may be the same crop.

NOTE: No more than 60 subclasses can be resident on the Fields Data Base at any time for a single segment.

B.1.2 LACIE FIELD CODES

Part C of figure B-1 is the LACIE Field Codes that are converted from the Analyst Field Codes shown in Part B of figure B-1. The descriptions of LACIE Field Codes are given in detail below. The LACIE codes must be used in all automatic processing.

B.1.2.1 Field Name

The field name is composed of three characters. The first character is an alphabetic corresponding to the field category code

for training and test fields. The codes are as follows:

Code	Category
W	Wheat
N	Other
X	Unidentifiable
P	Test

The other field types have the following alphabetics as the first character of their field names:

DO (Code D)
DU (Code U)

During automatic processing, the segment identifier (see section B.1.2.3.2) is added as the seventh character of the field name.

B.1.2.2 Field Type

A single character alphabetic designation with the following options is used.

Code	Type
T	Training
P	Test
D	DO
U	DU

B.1.2.3 Field Subclass Name

A six-character alphanumeric designation comprised of the category, segment identifier (*), and subclass qualifier is used. Training fields and test fields have subclass names. The DO and DU field types do not have subclass names. The six subclass characters have the following significance.

```
X          X                    XX                  XX
↑          ↑                    ↑                   ↑Subclass
|          |                    |                   |Qualifier
|          |                    |Class Identifier
|          |Segment Identifier
|Category
```

B.1.2.3.1 Category

A one-character alphabetic with the following options is used.

Category	Code
Wheat	W
Nonwheat	N
Unidentifiable	X
Test	P

B.1.2.3.2 Segment Identifier

This is encoded as an * and is automatically converted to the position of the segment in the merged image at the time of classification. When recalling this subclass name for automatic processing, the appropriate digit (0-3), must be used as the second character. In the normal situation where only one segment is used, the digit will always be 0. Only in cases of signature extension will this vary.

B.1.2.3.3 Class Identifier

A two-character alphabetic is used:

Class Identifiers	Codes
Winter wheat	WW
Spring wheat	SW
Small grain	GR
Winter small grain	WG
Spring small grain	SG
Agriculture	AA
Nonagriculture	NW

Class Identifiers	Codes
Unidentifiable	XX
Test	TS

B.1.2.3.4 Subclass Qualifier

The numerics used to describe the subclass qualifiers are given below for all categories except P. A two-character numeric, designating fields with similar signatures, is used. For example, two tree-covered areas might have the subclass identifier of "Ø1" but a city would be "Ø2". Each class has the potential of having subclass qualifiers from Ø1 to 50. (See the note in section B.1.1.2 for system restrictions.)

B.1.2.3.5 Test Category

The subclass qualifier for all test fields is 00. The DO and DU categories have no subclass identifiers.

NOTE: When the Winter Wheat (WW) field code is used to identify a signature or group of signatures within a particular segment/ acquisition, the only other wheat or grain classes that may be coded are Spring Wheat (SW) or Spring Grain (SG). When the Spring Wheat (SW) code is used, only the Winter Wheat (WW) or Winter Grain (WG) codes should be coded within the same segment. The Small Grain (GR) field code is used by itself, and no other code should be used with it. The classes that can be correctly used with the Winter Small Grain (WG) are Spring Wheat (SW) or Spring Small Grain (SG). When the Spring Small Grain (SG) is coded in a segment, only the Winter Wheat (WW) or Winter Grain (WG) codes may be used.

The following table illustrates the correct uses of multiple classes in a segment/acquisition.

If this class is used	One of these additional wheat classes may be used for the same segment/acquisition
WW	SW or SG
SW	WW or WG
GR	None
WG	SW or SG
SG	WW or WG

B.2 CHANGE CODE DESIGNATIONS

Change Codes will be used on the AI Temporal Crop Interpretation Form. Table B-2 gives the change code numbers and designations.

B.2.1 CHANGE CODE DESCRIPTIONS

The procedures for combining and splitting fields, adding new fields, changing the crop classification, and dropping a field are discussed in the following paragraphs.

B.2.1.1 No Change - Code A

No change has occurred in the field name, type, subclass, or the corner coordinates.

B.2.1.2 Crop Class Change - Code B

When a change in the crop class designation is necessary, the field number will not change. For example, a field may have been identified as spring wheat on the initial interpretation, but was later identified as barley on the revised interpretation.

B.2.1.3 Coordinate Change - Code C

When the size or shape of a field changes from one acquisition to the next, the coordinates must be reread and recorded. When there are no changes, the interpreter will not rerecord the

TABLE B-2. — CHANGE CODE NUMBERS AND DESIGNATIONS

Code	Subject
A	No Change
B	Crop Class Change
C	Coordinate Change
D	Crop Class Change and Coordinate Change
E	Added Field
F	Dropped Field
G	Split Field
H	Split Field and Crop Class Change
I	Combined Field
J	Combined Field and Crop Class Change
K	Obscured Field

coordinates, but will indicate on the Temporal Crop Interpretation Form that no change has occurred (Code A).

B.2.1.4 Crop Class Change and Coordinate Change - Code D

This code indicates a change in both the crop designation and corner coordinates of a field.

B.2.1.5 Added Fields - Code E

The Analyst may add training fields that are identified during the revised interpretation. These fields must be delineated, assigned a field name, recorded on appropriate forms, and the corner coordinates read. New fields created by "splitting" fields would also use this change code.

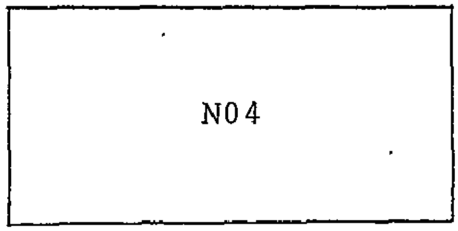
B.2.1.6 Dropped Field - Code F

A field no longer belonging to a crop class identified by the Analyst may be dropped. When this occurs, the field number will not be used again unless the analyst decides to recall this field to the same or a different class. A field may also be dropped if a better training field is identified. When a field is dropped, this will be recorded on the form, and thereafter, the information will be omitted.

B.2.1.7 Split Field - Code G

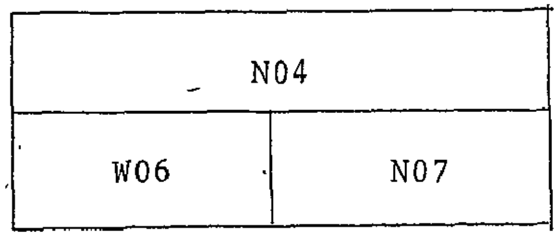
If a field is split into two or more fields with no change in the crop class, the new delineations will be made on the overlay. The original field number will be retained for one of the fields and new numbers will be assigned to the others. Corner coordinates must be read for these fields.

In the following example, the first field number (N04) is retained and two new numbers (W06 and N07) are assigned. Code G is assigned to field N04 and Code E (Added Field) to W06 and W07.



Initial Interpretation

May 22



Revised Interpretation

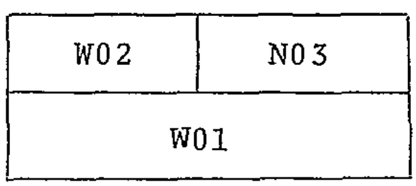
June 26

B.2.1.8 Split Field and Crop Class Change – Code H

When a field is split and the crop class changes in the portion that retains the original number, the corner coordinates must be read for these fields.

B.2.1.9 Combined Field – Code I

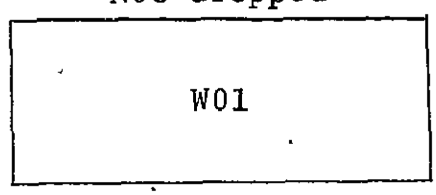
Should two or more fields now appear to be one large field, the new boundary will be delineated on the primary overlay. When this occurs, the smaller field number will be retained and all others will be dropped. This situation may occur during a revised interpretation. New corner coordinates must be read. An example of this is as follows.



Initial Interpretation

May 5

Field numbers W02 and N03 dropped



Revised Interpretation

June 11

B.2.1.10 Combined Field and Crop Class Change – Code J

This code is used when a field is combined with another field and the crop class changes from the previous dates designated. Corner coordinates must be reread.

B.2.1.11 Obscured Field - Code K

When clouds obscure a field or a system malfunction precludes the interpretation of the entire field, this code may be used.

APPENDIX C

EQUIPMENT

C.1 IMAGE INTERPRETATION EQUIPMENT

The Analyst is issued the following equipment: Portable light table for desk or Richards 3040 light table; 5X to 10X tube magnifier with reticle attachments; 2X to 5X magnifier on stand; rapidograph pen with a double or triple zero point; plastic triangle; and white gloves.

C.2 CONSOLE OPERATIONS

This section discusses the manual operation of the Building 17 consoles. The equipment in Building 30 is quite similar but not identical. For further explanation, the *LACIE User's Guide* and the *ERIPS Training Manual* may be consulted.

The scroll keys are located in the upper right hand corner of the key board. The labels for the gray screen are as follows:

- GH - Hold scroll
- GN - normal scroll
- GF - fast scroll

The labels for the color screen are as follows:

- CH - hold scroll
- CN - normal scroll
- CF - fast scroll

Both toggle switches are on the keyboard. The front toggle switch is for the gray level image screen and the conversational screen. The rear toggle is for the color screen. The switch for the image screen is toggled in the direction of the screen where the input is to be made. For example, if an input is to be made on the right-hand screen, then it should be toggled right.

The hard-copy machine does not produce the copy instantly. One machine is nominally attached to each console. However, in the event of malfunction of one of the machines, both consoles may use the same hard-copy machine. Care should be taken not to back up the copies. Both gray level screens can be copied. If the background is light, the screen is copied as is. If the background is dark, the image is reversed for the copy. Paper for the copy machine is kept in Building 30.

The conversational and image screen can be switched. This has no effect on the number of the image screen. The button is on the panel above the keyboard.

The gray level for an address can be changed. The enable button is released and the modify button enabled. The gray option is selected. The address button is then enabled and the two digit address is entered, then the gray level button is pushed and the desired gray level is entered. The addresses and gray levels are entered in this manner until all have been completed. To obtain the original values, the button marked RESTORE ORIGINAL TABLE is pushed along with the enable button. The modify key is then released.

APPENDIX D

CORNER COORDINATE RECORDING, REFORMATTING AND EDITING PROCEDURES

D.1 CORNER COORDINATE RECORDING (USING THE H. DELL FOSTER RSS-4MGT-2)

1. Load card punch machine with white cards.
2. Turn card punch machine on.
3. Turn Dell Foster Digitizer on.
4. Turn light table on.
5. Enable the REMOTE KEYBOARD by placing the KBD ENABLE switch to the up position. The red light will be on in this position.
6. Check to see that the REMOTE KEYBOARD is functioning correctly: Punch one or two test cards using the REMOTE KEYBOARD.
7. Place the digitizer mode switch to the X,Y-AXIS position.
8. Set the SCALE FACTORS to X-00551 and Y-00412.
9. Accurately square the imagery along the X-axis and the Y-axis: Move the cursor pip along an X-axis grid line then a Y-axis grid line and adjust the imagery so that the cursor pip remains directly over the grid line as the cursor is moved along a particular axis.
10. Zero the digitizer to a reference or starting point: Place the cursor pip directly over the 10X, 10Y grid intersection, press the CLEAR button to clear the digitizer then insert the coordinates 11X, 11Y into the digitizer.
11. Place a + sign in the digitizer for the X-axis and a - sign for the Y-axis.

PRECEDING PAGE BLANK NOT FILMED

12. Retest imagery alignment: Move the cursor to pixel 190, line 110. Check the digitizer readings. If the digitizer readings do not read 190X and 110Y, the imagery is not squared to the X and Y axis, or the cursor pip was not accurately placed over the 10X, 10Y grid intersection durint the zeroing in process.

13. Using the REMOTE KEYBOARD, type in the 4-digit segment number, the biophase acquisition date, and the Analyst name, then press the RETURN key.

14. Starting with Field No. 1, type in the three-character FIELD NAME, and if applicable, the FIELD SUBCLASS code and OPTION code 'B' from the AI Temporal Crop Interpretation Form (e.g. W01 or W01 W1 B); then press the RETURN key.

    If the SUBCLASS is not applicable, type in only the FIELD NAME and press the RETURN key, e.g., D11. See figure D-1.

15. Record field coordinates for Field No. 1. Place the cursor pip over each inked coordinate dot starting with the upper left coordinate and progressing in a clockwise direction. Record each coordinate by depressing the "RECORD" pedal located on the floor.

16. Repeat the two preceding steps for all succeeding fields.

17. After recording the last coordinate of the last field, check that the final coordinate on the imagery matches the digitizer readings.

18. Type //END. PRESS THE RECORD BUTTON.

19. If any fields are to be deleted, type in DEL and the FIELD NAME of the fields to be deleted, e.g., DEL W01 N07 D11.

20. Remove cards and check for errors.

```
1076 1 312 5 JOHNSON KILLEN
P01
+000002-000002
+000009-000002
+000010-000004
+000002-000004
N02 A1
+000036-000008
+000042-000006
+000045-000010
+000035-000012
N03 A2
+000045-000024
+000049-000023
+000053-000033
+000049-000033
N04 A3
+000034-000031
+000045-000029
+000046-000033
+000035-000035
W05 W1
+000001-000040
+000006-000040
+000007-000045
+000002-000048
W06 W1
+000006-000060
+000012-000059
+000016-000067
+000009-000067
W07 W2
+000038-000103
+000042-000100
+000044-000106
+000040-000107
W08 W2
+000081-000062
+000097-000059
+000097-000064
+000084-000069
W09 W2
+000082-000040
+000088-000040
+000089-000046
+000084-000049
D10
+000052-000001
+000130-000001
+000137-000020
+000118-000033
+000081-000039
+000081-000032
+000063-000032
N11 N7
+000114-000008
+000129-000002
+000130-000008
+000115-000011
D12
+000132-000002
+000195-000002
+000195-000117
+000182-000117
+000155-000050
+000120-000034
+000139-000020
N13 A1
+000128-000066
+000132-000066
+000131-000069
+000128-000069
W14 W1
+000076-000066
+000080-000064
+000085-000078
+000081-000078
N15 A2
+000092-000090
+000097-000090
+000099-000098
+000094-000099
P16
+000087-000102
+000069-000104
N17 N7
+000093-000112
+000103-000112
+000104-000116
+000094-000116
N18 N7
+000142-000084
+000152-000084
+000152-000091
+000144-000091
P19
+000029-000048
+000037-000046
+000043-000064
+000034-000064
P20
+000152-000114
+000156-000114
+000156-000117
+000152-000117
W21 W1
+000171-000113
+000176-000113
+000177-000117
+000172-000117
N22 A3
+000051-000036
+000062-000036
+000063-000040
+000053-000041
N23 A4
+000065-000035
+000076-000034
+000077-000038
+000067-000039
+000000-000000
// END
```

Figure D-1. — H. Dell Foster card format.

21. Secure the deck with card protectors and bind with a rubber band. This card deck is called the "DEAF" deck.

22. Proceed to the LARS Terminal.

D.2 FIELDS DATA CONVERSION AND EDITING ON THE LARS TERMINAL

1. Activate the LARS terminal by pressing the "ATTN" key.
2. Type in ID: L jsc602 (CR)
3. Type in PASSWORD: AIO (CR)
4. Type in NAME: smith (CR)
5. After "CP" type in i cms (CR)
6. Type in toad (CR). This initiates TOAD COMMANDER.
7. Type info (CR) if operation instructions and commands are desired.
8. Transmit DEAF deck:
   a. Place cards in card reader hopper with ID card first.
   b. Turn reader mode knob to "TSM-TRSP".
   c. Press "END OF FILE" button. Check that the END OF FILE light comes on.
   d. Press "START" button until transmission begins.
   e. When beep sounds, reader is through transmitting.
   f. Turn mode knob to "OFF LINE".
   g. Remove cards from reader.
   h. Press "NPRO" to flush cards through reader.
9. Type in start. (Start Pha-1 or Pha-2 or no map if no map is desired in output).

The following will occur:

a. With a DEAF deck free of errors detectable by the TOAD program.

(1) Indicate YES/NO whether segment class card is required.

(2) Beep will sound signaling that printer is ready to print.

(3) Turn mode knob to "TSM-TRSP".

(4) Press "START" button on printer.

(5) Products produced on printer at this point are as follows:

(a) MAP printout. See figure D-2.

(b) AI-DPA Summary (by field type). See figure D-3.

(c) FLAP printout in LACIE format. See figure D-4.

(d) AI/DPA Summary (Fields listed in numerical order.) See figure D-5.

b. With a DEAF deck containing errors detectable by the TOAD program:

(1) Type in correction(s) (CR).

(2) Processing will continue with the corrections incorporated in the final products.

(3) Indicate YES/NO whether segment class card is required.

(4) Beep will sound, signaling that printer is ready to print.

(5) Turn mode knob to "TSM-TRSP."

(6) Press "START" button on printer.

c. Products produced at this point are as follows:

(1) MAP printout

(2) AI/DPA summary

(3) FLAP printout in LACIE format

(4) AI/DPA summary

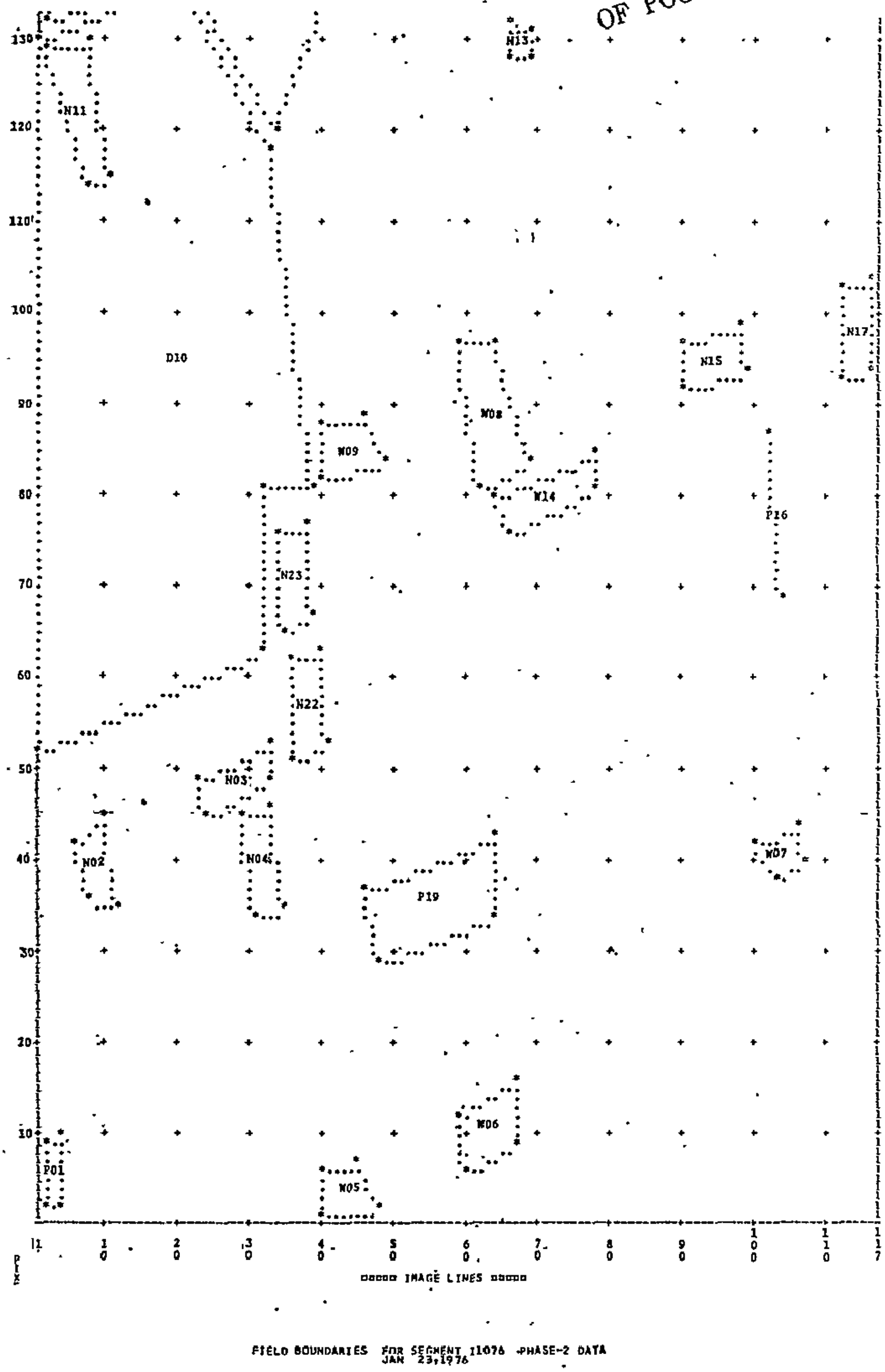


Figure D-2. — LARS map printout.

DPA/AI PHASE-2 FIELDS DATA BASE INFO FOR SEGMENT I1076

SEG. ID = 1076-1-312-5-JOHNSON -KILLEN

JAN 23,1976

CAT	APRIORI PROB.	THESHOLD
W	100	1.0
N	100	1.0
X	100	1.0
S	100	1.0
O	100	1.0
P	100	1.0

17 TRAINING

7 WHEAT POTENTIAL

FIELDS	CLASS NAME			
W05	W*WW01	WINTER WHEAT	TYPE	1
W06	W*WW01	WINTER WHEAT	TYPE	1
W07	W*WW02	WINTER WHEAT	TYPE	2
W08	W*WW02	WINTER WHEAT	TYPE	2
W09	W*WW02	WINTER WHEAT	TYPE	2
W14	W*WW01	WINTER WHEAT	TYPE	1
W21	W*WW01	WINTER WHEAT	TYPE	1

7 NON-WHEAT

FIELDS	CLASS NAME			
N02	N*AA01	AGRICULTURE	TYPE	1
N03	N*AA02	AGRICULTURE	TYPE	2
N04	N*AA03	AGRICULTURE	TYPE	3
N13	N*AA01	AGRICULTURE	TYPE	1
N15	N*AA02	AGRICULTURE	TYPE	2
N22	N*AA03	AGRICULTURE	TYPE	3
N23	N*AA04	AGRICULTURE	TYPE	4

3 NON-WHEAT

FIELDS	CLASS NAME			
N11	N*NN07	NON-AGRICULTURE	TYPE	7
N17	N*NN07	NON-AGRICULTURE	TYPE	7
N18	N*NN07	NON-AGRICULTURE	TYPE	7

4 TEST

FIELDS	CLASS NAME
P01	P*TS00
P16	P*TS00
P19	P*TS00
P20	P*TS00

2 DESIGNATED OTHER

FIELDS
D10
D12

Figure D-3. — LARS AI-DPA Summary by field type.

FLAP FILE CONTENTS PHASE-2 CODING

```
USER 1076-1-312-5-JOHNSON -KILLEN                                  1076
SEGSTART ID=I1076                                                  1076   1
SEGMENT CLASS=W*1076                                               1076   2
DEFLT CAT=W AP=100 TH=1.0                                          1076   3
DEFLT CAT=N AP=100 TH=1.0                                          1076   4
DEFLT CAT=X AP=100 TH=1.0                                          1076   5
DEFLT CAT=S AP=100 TH=1.0                                          1076   6
DEFLT CAT=D AP=100 TH=1.0                                          1076   7
DEFLT CAT=P AP=100 TH=1.0                                          1076   8
FLDSTART NAME=W05                                                  1076   9
FIELD  TYPE=T  CLASS=W*WW01                                        1076  10
FIELD PIXEL01=1   LINE01=40  PIXEL02=6   LINE02=40                 1076  11
FIELD PIXEL03=7   LINE03=45  PIXEL04=2   LINE04=48                 1076  12
FLDEND                                                             1076  13
FLDSTART NAME=W06                                                  1076  14
FIELD  TYPE=T  CLASS=W*WW01                                        1076  15
FIELD PIXEL01=6   LINE01=60  PIXEL02=12  LINE02=59                 1076  16
FIELD PIXEL03=16  LINE03=67  PIXEL04=9   LINE04=67                 1076  17
FLDEND                                                             1076  18
FLDSTART NAME=W07                                                  1076  19
FIELD  TYPE=T  CLASS=W*WW02                                        1076  20
FIELD PIXEL01=38  LINE01=103 PIXEL02=42  LINE02=100                1076  21
FIELD PIXEL03=44  LINE03=106 PIXEL04=40  LINE04=107                1076  22
FLDEND                                                             1076  23
FLDSTART NAME=W08                                                  1076  24
FIELD  TYPE=T  CLASS=W*WW02                                        1076  25
FIELD PIXEL01=81  LINE01=62  PIXEL02=97  LINE02=59                 1076  26
FIELD PIXEL03=97  LINE03=64  PIXEL04=84  LINE04=69                 1076  27
FLDEND                                                             1076  28
FLDSTART NAME=W09                                                  1076  29
FIELD  TYPE=T  CLASS=W*WW02                                        1076  30
FIELD PIXEL01=82  LINE01=40  PIXEL02=88  LINE02=40                 1076  31
FIELD PIXEL03=89  LINE03=46  PIXEL04=84  LINE04=49                 1076  32
FLDEND                                                             1076  33
FLDSTART NAME=W14                                                  1076  34
FIELD  TYPE=T  CLASS=W*WW01                                        1076  35
FIELD PIXEL01=76  LINE01=66  PIXEL02=80  LINE02=64                 1076  36
FIELD PIXEL03=85  LINE03=78  PIXEL04=81  LINE04=78                 1076  37
FLDEND                                                             1076  38
FLDSTART NAME=W21                                                  1076  39
FIELD  TYPE=T  CLASS=W*WW01                                        1076  40
FIELD PIXEL01=171 LINE01=113 PIXEL02=176 LINE02=113                1076  41
FIELD PIXEL03=177 LINE03=117 PIXEL04=172 LINE04=117                1076  42
FLDEND                                                             1076  43
FLDSTART NAME=N02                                                  1076  44
FIELD  TYPE=T  CLASS=N*AA01                                        1076  45
FIELD PIXEL01=36  LINE01=8   PIXEL02=42  LINE02=6                  1076  46
FIELD PIXEL03=45  LINE03=10  PIXEL04=35  LINE04=12                 1076  47
FLDEND                                                             1076  48
FLDSTART NAME=N03                                                  1076  49
FIELD  TYPE=T  CLASS=N*AA02                                        1076  50
FIELD PIXEL01=45  LINE01=24  PIXEL02=49  LINE02=23                 1076  51
FIELD PIXEL03=53  LINE03=33  PIXEL04=49  LINE04=33                 1076  52
FLDEND                                                             1076  53
FLDSTART NAME=N04                                                  1076  54
FIELD  TYPE=T  CLASS=N*AA03                                        1076  55
FIELD PIXEL01=34  LINE01=31  PIXEL02=45  LINE02=29                 1076  56
FIELD PIXEL03=46  LINE03=33  PIXEL04=35  LINE04=35                 1076  57
FLDEND                                                             1076  58
FLDSTART NAME=N13                                                  1076  59
FIELD  TYPE=T  CLASS=N*AA01                                        1076  60
FIELD PIXEL01=128 LINE01=66  PIXEL02=132 LINE02=66                 1076  61
FIELD PIXEL03=131 LINE03=69  PIXEL04=128 LINE04=69                 1076  62
FLDEND                                                             1076  63
FLDSTART NAME=N15                                                  1076  64
FIELD  TYPE=T  CLASS=N*AA02                                        1076  65
FIELD PIXEL01=92  LINE01=90  PIXEL02=97  LINE02=90                 1076  66
FIELD PIXEL03=99  LINE03=98  PIXEL04=94  LINE04=99                 1076  67
FLDEND                                                             1076  68
FLDSTART NAME=N22                                                  1076  69
FIELD  TYPE=T  CLASS=N*AA03                                        1076  70
FIELD PIXEL01=51  LINE01=36  PIXEL02=62  LINE02=36                 1076  71
FIELD PIXEL03=63  LINE03=40  PIXEL04=53  LINE04=41                 1076  72
FLDEND                                                             1076  73
FLDSTART NAME=N23                                                  1076  74
FIELD  TYPE=T  CLASS=N*AA04                                        1076  75
FIELD PIXEL01=65  LINE01=35  PIXEL02=76  LINE02=34                 1076  76
FIELD PIXEL03=77  LINE03=38  PIXEL04=67  LINE04=39                 1076  77
FLDEND                                                             1076  78
FLDSTART NAME=N11                                                  1076  79
FIELD  TYPE=T  CLASS=N*NN07                                        1076  80
FIELD PIXEL01=114 LINE01=8   PIXEL02=129 LINE02=2                  1076  81
FIELD PIXEL03=130 LINE03=8   PIXEL04=115 LINE04=11                 1076  82
FLDEND                                                             1076  83
FLDSTART NAME=N17                                                  1076  84
FIELD  TYPE=T  CLASS=N*NN07                                        1076  85
FIELD PIXEL01=93  LINE01=112 PIXEL02=103 LINE02=112                1076  86
FIELD PIXEL03=104 LINE03=116 PIXEL04=94  LINE04=116                1076  87
FLDEND                                                             1076  88
FLDSTART NAME=N18                                                  1076  89
FIELD  TYPE=T  CLASS=N*NN07                                        1076  90
FIELD PIXEL01=142 LINE01=84  PIXEL02=152 LINE02=84                 1076  91
FIELD PIXEL03=152 LINE03=91  PIXEL04=144 LINE04=91                 1076  92
FLDEND                                                             1076  93
FLDSTART NAME=P01                                                  1076  94
FIELD  TYPE=P  CLASS=P*TS00                                        1076  95
FIELD PIXEL01=7   LINE01=2   PIXEL02=9   LINE02=2                  1076  96
FIELD PIXEL03=10  LINE03=4   PIXEL04=2   LINE04=4                  1076  97
FLDEND                                                             1076  98
FLDSTART NAME=P16                                                  1076  99
FIELD  TYPE=P  CLASS=P*TS00                                        1076100
FIELD PIXEL01=87  LINE01=102 PIXEL02=69  LINE02=104                1076101
FLDEND                                                             1076102
FLDSTART NAME=P19                                                  1076103
FIELD  TYPE=P  CLASS=P*TS00                                        1076104
FIELD PIXEL01=29  LINE01=48  PIXEL02=37  LINE02=46                 1076105
FIELD PIXEL03=43  LINE03=64  PIXEL04=34  LINE04=64                 1076106
FLDEND                                                             1076107
FLDSTART NAME=P20                                                  1076108
FIELD  TYPE=P  CLASS=P*TS00                                        1076109
FIELD PIXEL01=152 LINE01=114 PIXEL02=156 LINE02=114                1076110
FIELD PIXEL03=156 LINE03=117 PIXEL04=152 LINE04=117                1076111
FLDEND                                                             1076112
FLDSTART NAME=D10                                                  1076113
FIELD  TYPE=D                                                      1076114
FIELD PIXEL01=57  LINE01=1   PIXEL02=130 LINE02=1                  1076115
FIELD PIXEL03=137 LINE03=20  PIXEL04=118 LINE04=33                 1076116
FIELD PIXEL05=81  LINE05=39  PIXEL06=81  LINE06=37                 1076117
FIELD PIXEL07=63  LINE07=32                                        1076118
FLDEND                                                             1076119
FLDSTART NAME=D12                                                  1076120
FIELD  TYPE=D                                                      1076121
FIELD PIXEL01=132 LINE01=2   PIXEL02=195 LINE02=2                  1076122
FIELD PIXEL03=195 LINE03=117 PIXEL04=182 LINE04=117                1076123
FIELD PIXEL05=155 LINE05=50  PIXEL06=120 LINE06=34                 1076124
FIELD PIXEL07=139 LINE07=20                                        1076125
FLDEND                                                             1076126
SEGEND                                                             1076127
```

FLAP FILE FOR SEGMENT I1076 CONTAINS 127 CARDS

Figure D-4. — FLAP printout in LACIE format.

DPA/AI PHASE-2 FIELDS DATA BASE INFO FOR SEGMENT I1076

SEG. ID = 1076-1-312-5-JOHNSON -KILLEN

JAN 23,1976

AI/CODE		FIELD NAME	CLASS NAME	TYPE	NO.VERT.
P01		P01	P*TS00	P	4
N02	A1	N02	N*AA01	T	4
N03	A2	N03	N*AA02	T	4
N04	A3	N04	N*AA03	T	4
W05	W1	W05	W*WW01	T	4
W06	W1	W06	W*WW01	T	4
W07	W2	W07	W*WW02	T	4
W08	W2	W08	W*WW02	T	4
W09	W2	W09	W*WW02	T	4
D10		D10		D	7
N11	N7	N11	N*NN07	T	4
D12		D12		D	7
N13	A1	N13	N*AA01	T	4
W14	W1	W14	W*WW01	T	4
N15	A2	N15	N*AA02	T	4
P16		P16	P*TS00	P	2
N17	N7	N17	N*NN07	T	4
N18	N7	N18	N*NN07	T	4
P19		P19	P*TS00	P	4
P20		P20	P*TS00	P	4
W21	W1	W21	W*WW01	T	4
N22	A3	N22	N*AA03	T	4
N23	A4	N23	N*AA04	T	4

Figure D-5. — AI/DPA summary of fields listed in numerical order.

10. After the FLAP printout has been produced, edit any errors which the TOAD program is not programmed to detect. Editing may be conducted in either the DEAF FILE or the FLAP FILE:

a. To edit the DEAF FILE:

(1) Type in editdeaf (CR)

(2) Locate item requiring editing by typing 1 (space)/ item to be located/(CR), e.g., 1 /wo2/ (CR).

(3) The desired item will be located and printed automatically: WO2 W1.

(4) To change an item, for instance, W1 to W2, type c (space) /the original entry/ the corrected entry/ (CR), e.g., c /w1/w2/ (CR).

(5) After all errors have been corrected, type in file (CR).

(6) Type in start deaf (CR). This initiates processing of the new DEAF FILE. The processing will progress as in a normal run.

b. To edit the FLAP FILE:

(1) Type in edit (CR).

(2) Locate line requiring editing by typing 1 (space) and placing within the slash marks something unique that is easily located, e.g., if FLDSTART NAME=WO2 is the line to be located, type in 1 /=wo2/ (CR).

(3) For example, to change WO2, to NO2, type c /wO2/nO2/ (CR).

(4) To delete a line type in d (CR). To delete more than one line type d (space) and the number of line entries to be deleted, e.g., d 5 will delete the located line and the next four lines.

(5) Once a line has been located and another nearby line is to be located type u or n (up or next, as appropriate) to locate the desired line, e.g., to locate three lines up, type u 3 (CR).

(6) After all errors have been corrected, type in file (CR).

(7) Type in print (CR). This will produce a printout of the new FLAP FILE.

11. After a correct FLAP printout has been produced, punch the FLAP FILE information onto cards.

   a. Beep will sound.

   b. Place sample amount of blank blue cards into reader hopper.

   c. Turn mode knob to "PUNCH" position.

   d. Press "START" button on the reader.

   e. When beep sounds reader is through punching cards.

   f. Remove excess cards from hopper to press the "NPRO" button to flush cards through the reader.

   g. Remove punched cards from reader.

   h. Remove punched JSC 602 card and excess blank cards from the card deck.

12. Type in log (CR) to log off the terminal if no other segment is to be processed.

13. Obtain an off-line printout of the punched FLAP deck:

   a. Place FLAP deck into reader hopper.

   b. Leave mode knob in the "OFF LINE" position.

   c. Press "START" on the printer and "START" on the reader.

   d. When printer stops, remove cards from reader and press "NPRO".

14. Secure the FLAP deck with card protectors and bind with a rubber band.

15. With a magic marker, write the segment number, biophase, and acquisition date on the top edge of the FLAP deck.

16. Return JSC 602 ID card.

D.3 BATCH POST PROCESSING OF CAMS REPORT FICHE

a. Each day, as all LACIE segments are processed, a CAMS Report Tape is operated.

b. CAMS personnel locate the tape numbers and retrieve the tapes from Building 30 daily.

c. A card deck and JSC Form 588A are submitted with the tape to Building 12 for processing.

Below is an example deck for this job:

Card 1. @RUN, R/R KAT17D, 1666M-ASP2-C-TF4-L88371, 15, 200
K. Abotteen

Card 2. @MSG, N TAPES 014320/F

Card 3. @USE L., TF4-L88371 *LBS.

Card 4. @ASG, AX L.

Card 5. @ASG, T/F 1, 8C, 014320

Card 6. @ADD L. RUNSTM

Card 7. @FREE 1.

Card 8. @FIN

Card 1 contains the identification of the user. Cards 2 and 5 contain the tape number and must be changed accordingly. If more than one tape is to be processed (no more than two tapes are expected) then the following format should be used.

Card 1. Same

Card 2. Same but Tape #1

Card 2. Same but Tape #2

Card 3. Same

Card 4. Same

Card 5. Same but Tape #1

Card 6. Same

Card 7. Same

Card 5. Same but Tape #2

Card 6. Same

Card 7. Same

Card 8. Same

d. The tape is returned to Building 17 via courier where it is picked up by the Data Clerk, who distributes the products to the appropriate Analysts.

APPENDIX E

PROCESSOR CAPABILITIES

This section briefly addresses the capabilities of the LACIE system. It does not attempt to outline procedures for the use of each processor. These are addressed in the nominal run or rework guidelines. In general, these capabilities are used for special cases and rework situations and the usage is tailored for the requirements of the individual segment.

Accompanying each processor section is the batch card format that is applicable to it. The numbers refer to card type, not sequence.

For further explanation of the processors and batch run set-up, the *LACIE User's Guide* should be consulted.

E.1 COMPUTATION OF STATISTICS

This section describes briefly the types of statistics and principles of clustering and chaining which may be employed in the classification processing of a segment. Options available under various situations to the Analyst are also discussed.

E.1.1 NORMAL STATISTICS

The statistics processor computes statistics (means, standard deviation, and covariance matrices) for training fields. The default mode for the batch system is the computation of subclass and field statistics; therefore, no card input is required. Subclass statistics are the only statistics that are retrievable for other statistics manipulation.

Interactively, statistics may be deleted as well as combined. However, automatic deletion does not take place and the names given to output statistics may not duplicate existing subclass names. In the batch system, the statistics processor has the capability to combine classes by using the following card.

CARD 11: COMB=BBBBBB(AnAA,A,AnAAAA)

This card combines all statistics input into one subclass with the name BBBB. Only 15 input names are permitted.

E.1.2 CLUSTERING

Clustering is the process of grouping pixels according to some distance measure. The iterative clustering algorithm uses ℓ_1 distance measure for data grouping. The ℓ_1, distance between two pixels $\overline{X}_i$ and $\overline{X}_j$, is defined as

$$\ell_1(\overline{X}_i, \overline{X}_j) \triangleq \sum_{k=1}^{n} |X_{ik} - X_{jk}|$$

where n is the number of channels used in the clustering.

The ERIPS clustering capability can be used in a variety of ways. For instance, clustering may be used to test the homogeneity of supposedly homogeneous training fields. In cases where the underlying probability distributions for a particular class is multimodal, clustering may be used to attempt to isolate the individual modes or subclasses. To what extent this is successful depends on the separation between the modes of the distribution and on the specific parameters that are input to the clustering algorithm.

The iterative algorithm generates clusters by making a variable number of passes through the data (i.e., fields/classes to be clustered).

The following table shows the default parameters available in batch mode.

Parameters	Default 1	Default 2
Maximum Number of Clusters	10	20
Strip Generation Parameter	5	5
Select Iterative Algorithm Box	Yes	Yes
Percent	.8	.8
SEP	1	1
STDMAX	3.2	3.2
DLMIN	3.2	3.2
NMIN2	2* (# of chs)	2* (# of chs)
ITMAX	5	8
Split/Combine Sequence	SC	SC
Select STAT Box	Yes	Yes
Select SBCLS Statistics Vector Box	Yes	Yes

Interactively, the capability exists to input all the above

parameters, altering any value as well as the capability to default to either of the above sets of parameters.

CARD 5: CLCC=nn,nn,nn

This card indicates the channels to be used for all clustering and chaining in this run. Statistics are computed and stored for all channels even though only the channels appearing on the CLCC are used for clustering and/or chaining. Only one of these cards may be present in a card group.

Example: CLCC=05,03,1

The following two card types relate to clustering alone and use the above table of parameters.

CARD 6: CLU1=BBBB(ANAAAA,ANAA)

This card initiates all iterative clustering processors using Default 1 parameters.

Example CLU1=W1W1(W1WW)

Output subclasses would be W1W101-W1W110. The statistics for all subclasses in W1WW would be deleted.

CARD 7: CLU2=BBBB(AnAAAA,ANAA)

This card has the same function as Card 6, but uses Default 2 parameters.

Example: CLU2=N*NN(N1A1,N1A2,N1N1)

The following cards are used for a combination of chaining and clustering. The chaining algorithm is described in section E.1.3.

CARD 8: CCH1=BBBB(AnAAAA,AnAA)

This card causes clustering using Default 1 parameters as in Card 6. However, before statistics are stored, one pass through the Bhattacharyya chaining algorithm is made using a threshold of 0.5. The statistics, which are stored, are the results of chaining the cluster output. The cluster map generated will

also reflect chaining results rather than cluster results.

Example: CCH1=W1WC(W1WW)

<u>CARD 9</u>: CCH2=BBBB(AnAAAA,AnAA)

This card causes clustering to be executed using Default 2 parameters (as in Card 7) and then pass through the chaining processor (as in Card 8).

Example: CCH2=N*NN(N1NW)

For a batch cluster, two situations exist: No fields reside on the Fields Data Base for this segment, and fields which should be retained are present on the data base for this segment.

In the first situation, the Fields Data Base must be initialized using a dummy field. Then a batch job is submitted for the cluster run. Below are example jobs for this situation.

Example A. Fields Data Base Initialization

```
USER 1-3315-3635-CWH-04-C977A
SEGSTART ID = I1052
SEGMENT CLASS = W* 1052
FLDSTART NAME = W*0
FIELD CLASS = W*WWWW TYPE = T LINE01 = 1 PIXEL01 = 1
              LINE02 = 1 PIXEL02 = 10
FLDEND
SEGEND
```

Example B. Cluster Job for Single Acquisition, Dummy Field

```
NAME = 1-3315-3635-CWH-07-C978
RSEG = 1052
CSEG = (75331)
CLU2 = CLU2(W)
```

Example C. Cluster Job for Multiple Acquisitions, Dummy Field

NAME = 1-3315-3635-CWH-07-C978
RSEG = 1052
CSEG = (75291, 75331)
CLU2 = CLU2(W)

For the second situation, the fields currently on the data base may be used as starting vectors for the clustering algorithm. In this case, no Fields Data Base job need be run. The two examples below show the format for this situation.

Example D. Cluster Job with Single Acquisition, Previous Fields

NAME = 1-3315-3635-CWH-07-C978
RSEG = 1052
CSEG = (75331)
CLU2 = CLU2 (W,N)

Example E. Cluster Job with Multiple Acquisitions, Previous Fields

NAME = 1-3315-3635-CWH-07-C978
RSEG = 1052
CSEG = (75291, 75331)
CLU2 = CLU2 (N,W)

For further explanation of the batch card formats and clustering algorithm, consult the *LACIE User's Guide*.

For interactive clustering, the Analyst may select the parameters which he feels would best separate the signatures on the segment. The main difference between the batch and interactive clustering is the starting vectors. Batch processing utilizes training fields as starting vectors, while in interactive clustering

self-starting vectors are used. For details on console operation and interactive clustering, consult the *LACIE User's Guide.*

The outputs of both methods are a Color Cluster Map and a PFC Product 9. However, the clustering output for the same acquisition may not be identical using the above methods.

Example F.

Interactive Example

a. Sign on
b. Create merged image with desired acquisition dates
c. Enter merged image name or initial PR menu
d. Define a field which encompasses the entire segment
e. Select clustering
f. Select all channels
g. Enter first letter of subclass name or field defined in d
h. Enter clustering parameters

Maximum number of clusters	16
Strip generation parameter	5
Iterative	Check box
Percent	.8
SEP	1.0
STDMAX	3.2
DLMIN	3.2
NMIN2	50
ITMAX	8

Check self-starting vectors box

i. Hard-copy cluster report menu
j. Select map generation box
k. Check clustering image tape
l. Default remaining map menus

E.1.3 CHAINING

Chaining is a process of signature manipulation which utilizes distance measure to combine similar subclasses or clusters together. The distance measure used in the LACIE chaining is the complement Bhattacharyya distance defined as

$$H_{ij} = 1 - H'_{ij} \ ,$$

where H'_{ij} is the Bhattacharyya distance for signatures i and j. The function H'_{ij} which characterizes the similarity between signatures i and j takes a simple form.

$$H'_{ij} = \mathrm{Exp}\left\{ -\frac{1}{4}\left(M_i - M_j\right)^T \left(\Sigma_i + \Sigma_j\right)^{-1} \left(M_i - M_j\right) - \frac{1}{2}\ln \frac{\left|\frac{\Sigma_i + \Sigma_j}{2}\right|}{|\Sigma_i|^{\frac{1}{2}}|\Sigma_j|^{\frac{1}{2}}} \right\} \text{ or}$$

$$\left|\frac{1}{2}\left(\Sigma_i + \Sigma_j\right)\right|$$

For a normal distribution, where Σ_1, Σ_j, M_1 and M_j are the covariance matrices and mean vectors for signatures i and j, respectively. The value of H'_{ij} varies from zero to unit. This implies that the complement will range from 0 to 1 also. Clearly, under the criterion of the complement Bhattacharyya distance function, the more similar the two signatures are, the smaller the value is.

While the principle of chaining is rather simple, the actual computer programming involves a complicated recursive technique. The principle is that two groups of signatures, X and Y, are close enough to be merged into the same group if and only if the following tests are met:

- All H_{ij} for i contained in X and j contained in Y are less than or equal to a predetermined chaining threshold.

- $\overline{H}_{xy}$ is less than or equal to all possible $\overline{H}_{xu}$ and $\overline{H}_{uy}$ where u denotes a group of signatures other than X and Y.

The bar stands for the average, i.e., the sum of all pair distances between the two groups involved divided by the number of pairs.

Chaining in the LACIE system can be accessed from pattern recognition for chaining of subclasses or from clustering for cluster chaining. In the interactive mode, the chaining threshold can be manually input by the Analyst and the number of chaining cycles is not limited. However, in the batch mode the chaining threshold is set at 0.5 and only one chaining cycle is performed per access to the chaining processor. For other restrictions and features of the chaining functional capabilities, the user is referred to the *LACIE User's Guide*.

The primary purpose of performing chaining is to statistically manipulate the signatures to improve the statistics to be used for classification. Another application is to examine the uniformity of the subclasses. These objectives can be most effectively achieved by an interactive mode run because many chaining thresholds can be tested in the run and the results are obtained immediately. The following procedures may serve as a guide to perform the chaining process interactively.

- To upgrade subclass statistics, one can perform subclass chaining using various chaining thresholds and compute the statistics for each cycle until a satisfactory result is achieved.
- To examine the subclass uniformity, one can exercise subclass chaining. The uniformity is judged by whether similar subclasses are partitioned into the same subgroup.

- The Bhattacharyya distance is a function which decreases as the number of channels increases. Thus, the chaining threshold should increase with an increase in channel number. For example, a threshold of 0.5 may be too small to cause chaining for 16 channels but too big for 4 channels.

To perform chaining alone the following card is necessary.

<u>CARD 10</u>: BCHN=BBBB(AnAAAA,AnAAAA,AnAAA)

Bhattacharyya chaining will be executed with a threshold of 0.5. The subclasses input must be from the original subclass statistics table and will be replaced with the chained statistics named BBBB01-BBBB20.

To perform chaining, followed by clustering, for a particular subclass or group of subclasses, use the cards (Types 8,9) described in section E.1.2.

E.1.4 BATCH RESTRICTIONS

Statistics option cards are heirarchically equal and will be executed in the order in which they appear in the card group. The following naming restrictions apply to these cards:

- Since each statistics option deletes the statistics for input subclasses before storing outputs, input subclasses may not appear on more than one statistics option card. This includes implied input subclasses. For example BCHN=W1WC(W1WW) means that <u>all</u> subclasses in class W1WW would be deleted and replaced by chains W1WC01-W1WC20; and later input of CLU1=W1WC(W1WW01,W1WH) would be invalid.
- Once original statistics have been replaced by modified statistics, the new subclasses may not appear as input names. For example,

  BCHN=W1W1(W1WW)

  CLU1=W*WW(W1W1) would not be a valid combination.

- Output names must be exactly four non-blank characters, except for COMB where the output name must be six nonblank characters.
- Inputs names must have a numeric second character. This numeric must refer to the position in the merged image of the segment to which the categories, classes, or subclasses belong.
- The output name on each COMB cards must be unique, i.e., it may not be duplicated on any statistics option cards output name.
- A maximum of 20 input names per card is allowed.

E.2 MODIFICATION OF STATISTICS

The purpose of the modification of statistics is to allow for the statistical manipulation of categories, classes, subclasses, and/or the entire segment. In general, these capabilities were designed for use in signature extension. Specifically, the following options are available for the modification of statistics in both the batch and interactive processing modes:

- Mean Level Adjustment (MLA)
- Sun Angle Correction
- External Statistics

CARD 4: TSEG=nnnn(YYddd,yyddd)

This card identifies the training segment and its dates of acquisition to be used in classifying all sites on the RSEG card; 0-3 of these cards may be present. The order of TSEG cards should be noted since all field and subclass names will be altered to reflect the position of the segment in the merged image. The number of dates of acquisition must agree with the

number of dates of acquisition on CSEG cards. This card must precede the "MLAn" card, if included in the input batch deck.

Example: TSEG=1087(74289,74361).

E.2.1 MLA

Through use of the MLA option, LACIE will allow the adjustment of training signatures gathered on training segments, to provide better representation of recognition segment data. There can be a single MLA for a segment of a weighted MLA averaging over multiple corrections on more than one segment.

For MLA_i (the MLA relative to training segment i or, if i=v, the averaging of multiple adjustments), the subclass mean vector μ_A of training class A is modified according to the following formula:

$$\mu_A(\text{new}) = \mu_A(\text{current}) - \Delta\mu_{Ti}.$$

$$\Delta\mu_{Ti} = \mu_{MLA(Ti)} - \mu_{MLA(R)}$$

where

$\mu_{MLA(Ti)}$ = Mean Level Adjustment vector for the i^{th} training segment.

$\mu_{MLA(R)}$ = Mean Level Adjustment vector for the recognition segment.

for i = 0, 1, 2, or 3.

The keyword MLA on Card 15 will indicate that the categories, classes, or subclasses listed following the equals sign are to be corrected for mean level.

MLA0 = XXXXX,XXXXXX,XXXXXX,XXXXXX

MLA1 = XXXXX,XXXXXX,XXXXXX,XXXXXX

MLA2 = XXXXXX,XXXXXX,XXXXXX,XXXXXX

MLA3 = XXXXXX,XXXXXX,XXXXXX,XXXXXX

The MLA will be performed between training segment n and the recognition segment for those subclasses, classes, and categories appearing as input names. If MLAV appears, the weighted average of all MLA fields contained within all training segments will be used to perform the correction. If the user specifies "ALL" instead of the subclass list, all subclass statistics will be corrected. When the ALL option is specified, no other MLA cards should appear. A maximum of 60 names may be input on the group of MLA cards.

Interactively, select MLA on the statistics menu. Enter categories, classes, and subclasses to be corrected. The option AVG may be selected to give the weighted average of all MLA fields contained within all training segments. The MLA correction on categories, classes, or subclasses can be verified by examining the fields report and comparing values of previous field reports.

E.2.2 SUN ANGLE CORRECTION

LACIE allows for the correction of training signatures to account for the difference in the Sun angle of acquisition between recognition segment and training segment. Both the mean vector μ_C and the covariance matrix Γ_C of a subclass C will be modified according to the following formulas:

$$\mu_C(\text{new}) = A \cdot \mu_C(\text{old}) + B \cdot U$$

$$\Gamma_C(\text{new}) = A \cdot \Gamma_C(\text{old}) \cdot A,$$

where A and B are diagonal matrices, whose diagonal entries are obtained from the Sun Angle Table, and each coordinate of the vector U is 1.

CARD 14: SUNn=AAAAA,AAAAA

This card causes Sun angle correction to be made between training segment n and the recognition segment. Subclasses, classes, and categories appearing as inputs are to be corrected. If SUNn=ALL is input, it is the only SUN card allowed. Therefore, if corrections are to be made to all categories of all training segments, three cards must be input: SUN1=W1WW,N1NW

SUN2=W2WW,W2NW,X2XX

SUN3=W3WW,N3NW,X3XX

Sixty subclass names may be input. It should be noted that all classes must be input since the category refers to that category in ALL segments.

E.2.3 EXTERNAL STATISTICS

LACIE allows for signature extension through use of the external statistics retrievable either by checking appropriate boxes on the Signature Extension Menu, or through card input. Statistics tapes for retrieval may be created either in batch or interactive mode.

CARD 13: STAT=nnnnn

This card causes an external statistics tape to be written after all statistics manipulations have been completed. The tape number may be input or left blank.

Example: STAT=

CARD 12: EXTN=nnnnnn

This card causes an external statistics tape to be input just prior to fields retrieval. If this card is input no other statistics manipulations will take place. Card 12 and Card 13 may not both be present in the same card group.

Example: EXTN=12345

E.3 FEATURE SELECTION

The objective of the Feature Selection process is to reduce the dimensionality of the data while preserving separability to distinguish classes. In the LACIE system, the Feature Selection processor determines the best subset of channels according to the Bhattacharyya distance through the "without replacement" search procedure. Essentially, there are five optional paths one can choose from the feature selection processor. They are (1) computation of SEPARABILITY ONLY[1] for all available channels over all pairs of subclasses; (2) WITHOUT REPLACEMENT search for an optimal channel set of at least three channels which exceeds the SEPARABILITY THRESHOLD with respect to all channels; (3) computation of relative separability for the CHANNEL SUBSET INDICATED BELOW; (4) specification of CHANNEL SUBSELECTION to remove certain channels from further consideration in the search for an optimal subset coupled with separability only or (5) with "without replacement" search.

While all of the aforementioned options are available in the interactive mode, only the first three paths are applicable to the batch mode. Furthermore, in the interactive mode, the last two options, CHANNEL SUBSET INDICATED BELOW and CHANNEL SUBSELECTION, may not be selected simultaneously.

In batch mode, where Feature Selection is always executed, the separability threshold is set at 0.3 and cannot be overriden. The *a prioris* may be manipulated through the use of Card 18, discussed in section E.4.1.

[1]Capitalized words show the exact wording on the Feature Selection Menu.

Card 16 is used to alter the default sequence which is as follows:

- If the number of channels is less than or equal to 8, SEPARABILITY ONLY is computed.
- If the number of channels is greater than 8, WITHOUT REPLACEMENT is computed.

CARD 16: FSEL=ALL,CAT=A
FSEL=(nn,nn,nn),CAT=A
FSEL=WRO,CAT=A

This card has one of the three above forms. Only one FSEL card is allowed in a card group. "ALL" causes separability only to be executed; (nn,nn,nn) allows the user to input a specific channel subset which will be used for classification and for which a relative separability report will be output; WRO causes the "without replacement" option to be executed, a separability report to be generated, and the selected channels will be used in classification. These paths are mutually exclusive. For example, WRO and (nn,nn,nn) cannot both be executed in one card group. This card overrides all default feature selection. For example, if there are 16 channels available and FSEL=ALL is input, no reduction of the dimensions of data will be done and all channels will be used in classification. The CAT-A field is optional on all cards and specifies which category represents the category Other.

In the interactive mode, all options are available. The separability threshold may be manipulated and a prioris may be input. Contrary to the batch system, the execution of Feature Selection is not mandatory interactively.

E.4 CLASSIFICATION

The purpose of the batch mode classification is to classify all pixels of all training and test fields in the merged image, and to present tabular classification results through a detailed

classification summary report. The classifier will utilize the following density function in LACIE:

$$P_K(\bar{X}) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{q_k}{m_i n_k} (2\pi)^{-d/2} \cdot |\Gamma_{i,j}|^{-\frac{1}{2}} \cdot e^{y/2}$$

where $y = (\bar{X} - \mu_{ij})^T \cdot \Gamma_{ij}^{-1} \cdot (\bar{X} - \mu_{ij})$, and

d = Dimension of vector $\bar{X}$.

K = Category identifier.

i = Class identifier.

j = Subclass identifier.

n_k = Number of classes in category k.

m_i = Number of subclasses in class i.

q_i = A priori probability of category k.

Γ_{ij} = Subclass covariance matrix.

μ_{ij} = Subclass mean vector.

For ordinary pixels, the pixel will be assigned to the category K having the largest density function value. Then the classified pixels will be thresholded by using the algorithm described in section E.4.

The capability exists to classify either to the class or to the category level in both batch and interactive modes. The default is to classify to the category level.

<u>CARD 19</u>: CCAT=A,A,A,A
CCAT=ALL

This card indicates which categories are to be summed to the class level rather than the category level in classification.

Example: CCAT=N

No card input is needed for classification. It will automatically be performed.

Interactively, many options exist for the classification manipulations. Subclasses may be excluded, only test fields or training fields may be classified and so forth.

E.4.1 *A PRIORI*

A priori is the probability that an event will occur and is calculated prior to observing the result of one or more trials of the random process concerned. The values for the *a priori* probabilities for a given area can be developed from the historical data of the area. The *a priori* percent values for use in classification and feature selection are arrived at from the collection of category integers using the following procedure:

a. If the total of all category integers exceeds or equals 100, then categories with a zero integer are completely eliminated from the classification algorithm and the other categories are each assigned a percentage equal to their integer's portion of the category total.

b. If the total of all category integers is less than 100, then:

   (1) If no categories have zero for their integer, the categories' *a priori* values are not modified. This has the same effect as converting them to a percentage of the total.

   (2) If some categories have a zero integer, then each of them is assigned an equal portion of one one-hundreth of the difference between the total of the category integers and 100. This converts the integers to a decimal value.

c. If no integer is entered for a certain category, then the displayed default value becomes the integer for that category. Integers entered in positions without corresponding category are disregarded.

d. Each category *a priori* is then divided equally among its classes. Each resultant class *a priori* is then divided equally among its subclasses.

e. Default *a prioris* are taken from the Fields Data Base, but may be overriden in batch or interactive mode. To override the data base values in batch, Card 18 should be input.

f. The *a prioris* actually used will appear on the classification summary report.

g. Category names entered without integers are assigned a zero value as a default.

CARD 18: APRI=(A,nnn),(A,nnn)

This causes the *a prioris* on the data base to be overriden for those categories listed. The normalization rules apply as in interactive mode. It should be noted that if, through overrides, the sum of all categories *a priori* is > 100 and there are still categories which have 0 *a prioris*, those categories will have no pixels assigned to them.

E.4.2 THRESHOLDING

Classification will be performed to the category level in the LACIE system. The following steps will be used for thresholding the classified pixels:

a. Determine the category with the largest $P_k(\overline{X})$ for all k.

b. Compute the smallest $T_{k,ij} = (\overline{X} - \mu_{ij})T \cdot \Gamma_{ij}{}^{-1} (\overline{X} - \mu_{ij})$ for all i and j for this category.

c. Pixel $\overline{X}$ will be assigned to the class "threshold", if $T_{k,ij}\ t_k$, where t_k is the chi-square percentage point.

d. If $T_{k,ij} < t_k$, assign pixel $\overline{X}$ to the category k.

If a category is given a zero threshold value, no pixels of that category are thresholded out of classification results and, conversely, a category given a 100 threshold value will have almost all pixels thresholded out of that category.

In both the interactive mode and batch mode, threshold values are accepted in the form XXX.X and rounded down to the nearest one-half. A threshold value entered in the "ALL CATEG" blank will be used for all categories that do not have individual threshold values overriden.

CARD 17: THRS=(A,nnn.n),(A,nnn.n)

This denotes a threshold for the categories listed. These will override the data base values for this run only. These threshold values will be used only in classification. Values allowed are from 0.0 to 100.0.

APPENDIX F

GLOSSARY

Analyst Team – two Analysts, an Analyst Interpreter and a Data Processing Analyst, who do all automatic, interpretive, and evaluative processing of a sample segment. A Regional Analyst, on a consulting basis, is also a member of the team.

Biological Stage – an external morphological growth development, usually measured on a scale from 1-9.

Biophase – biostage

Biowindow – the dates, according to the nominal crop calendar, between which a particular growth stage is expected. Currently, the growing season is divided into 4 biowindows as defined by the CAMS Requirements Document.

Character Map – an alphanumeric representation of the assignment of pixels by the classification algorithm or the clustering algorithm.

Classifier Accuracy – a pixel within a training field is considered correctly or accurately classified if it is classified into the same class or category as the entire training field was labeled.

Classification Acquisitions – those acquisitions for which training fields are selected and/or are processed through the automatic system.

Classification Map – a black and white film product which reflects the spatial relationship of the pixels of wheat/nonwheat/unidentifiable/threshold as assigned by the classification algorithm.

Classification Summary — a detailed report reflecting the performance of each field defined. For each field, the number of pixels classified into and thresholded from each category, class and subclass is listed. For the entire segment, the number of pixels is altered to account for exclusion areas. A classification accuracy for each field is given on the subclass level and for the entire segment at the category level.

Classifier — the algorithm by which a pixel on a LACIE sample segment is assigned to a subclass, class and category. The term may also be used to describe the entire automatic data processing system.

Class Match — the subclass to which cluster statistics are closest based on a statistical distance measurement.

Cluster Map — a color film product which shows the spatial representation of the assignment of pixels to clusters based on the iterative clustering algorithm.

Color IR — false color infrared image produced from channels 1, 2 and 4. (Product 1)

Confusion — a nonwheat signature which on one acquisition appears the same as one of the wheat signatures in the scene.

Crop Calendar — a calendar depicting the period of the growth-development or biological stages of the major crop types within a specified region during a calendar year. The seedbed preparation, planting and harvest are included on the calendar.

Crop Calendar Adjustment — Identifying the correct phenological stage of wheat band on the Robinson Model. The model predicts the growth stage from the maximum and minimum daily temperature

and this is translated into plus or minus days from the nominal crop calendar.

Crop Class – a homogeneous or representative signature that is identified as being unique and different from the other signatures on the color infrared or other imagery.

Data Quality – physical or technical characteristics of the imagery.

Designated Unidentifiable – an area which has been obscured by natural phenomena such as clouds, haze, or snow.

Initial Interpretation – the first wheat and nonwheat training field selection for a particular segment.

Interpretable – an acquisition which is free of technical and physical difficulties so that a judgement can be made on the processibility of that acquisition.

Landsat Data Set – the film products produced on a particular acquisition of a sample segment for the purpose of identifying wheat.

Positive-Negative Imagery – a multiband color film composite produced by adding together positive and negative bands. (Product 2).

Primary Overlay – the overlay keyed to the color IR imagery; contains the final field boundary delineations.

Process – to estimate and forward to CAS the percentage wheat on a given acquisition of a sample segment.

Processible — an acquisition which has been determined to contain information which would enhance the wheat estimate for that sample segment.

Review Image — Landsat acquisitions that are interpreted for potential value as classification acquisitions.

Revised Interpretation — Selection of new or additional training fields for the purpose of improving the wheat estimate within the sample segment. In addition, re-estimating the wheat.

Screening Imagery — Landsat data set.

Summary of Summary — a computer printout which lists all the necessary percentages and is not broken down to the detail of the classification summary. It shows class and category percentages and accuracies, rather than subclass.

Test Field — a homogeneous, if possible, field defined, but not identified. The field is treated as a training field in automatic processing except that no statistics are generated for it.

Test Film — product used for a weekly quality assurance check on the PFC imagery.

Training Field — a homogeneous area which represents a crop class. This field is used for statistics computation during automatic processing.

Unidentifiable — the category for training fields for obscured areas.

Wheat Identification Aid — a crop identification manual that provides a comparison of different crops and their corresponding signatures on the Landsat B/W and color composites.

APPENDIX G

REFERENCES

1. *LACIE User's Guide*, vol. 1, IBM, April 1975.

2. *ERIPS Training Manual*, LEC-3708, June 1974.

3. *Wheat Identification Aid*

4. *LACIE Baseline Document*, vol. 1, IBM, July 1975.

5. *LACIE Requirements Document*, vol. 3, December 1975.

6. *CAMS Functional Flow Procedures*, to be published.